COMPUTATIONAL APPROACHES FOR ANALYZING SOCIAL SUPPORT IN ONLINE HEALTH COMMUNITIES

Hamed Khan Pour

Dissertation Prepared for the Degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

May 2018

APPROVED:

Bill Buckles, Major Professor
Cornelia Caragea, Co-Major Professor
Paul Tarau, Committee Member
Song Fu, Committee Member
Barrett Bryant, Chair of the Department of Computer Science and Engineering
Costas Tsatsoulis, Dean of the College of Engineering
Victor Prybutok, Dean of the Toulouse Graduate School

Khan Pour, Hamed. *Computational Approaches for Analyzing Social Support in Online Health Communities*. Doctor of Philosophy (Computer Science and Engineering), May 2018, 87 pp., 17 tables, 15 figures, 127 numbered references.

Online health communities (OHCs) have become a medium for patients to share their personal experiences and interact with peers on topics related to a disease, medication, side effects, and therapeutic processes. Many studies show that using OHCs regularly decreases mortality and improves patients' mental health. As a result of their benefits, OHCs are a popular place for patients to refer to, especially patients with a severe disease, and to receive emotional and informational support. The main reasons for developing OHCs are to present valid and high-quality information and to understand the mechanism of social support in changing patients' mental health. Given the purpose of OHC moderators for developing OHCs applications and the purpose of patients for using OHCs, there is no facility, feature, or sub-application in OHCs to satisfy patient and moderator goals. OHCs are only equipped with a primary search engine that is a keyword-based search tool. In other words, if a patient wants to obtain information about a side-effect, he/she needs to browse many threads in the hope that he/she can find several related comments. In the same way, OHC moderators cannot browse all information which is exchanged among patients to validate their accuracy. Thus, it is critical for OHCs to be equipped with computational tools which are supported by several sophisticated computational models that provide moderators and patients with the collection of messages that they need for making decisions or predictions. We present multiple computational models to alleviate the problem of OHCs in providing specific types of messages in response to the specific moderator and patient needs. Specifically, we focused on proposing computational models for the following tasks: identifying emotional support, which presents OHCs moderators, psychologists, and sociologists

with insightful views on the emotional states of individuals and groups, and identifying informational support, which provides patients with an efficient and effective tool for accessing the best-fit messages from a huge amount of patient posts to satisfy their information needs, as well as provides OHC moderators, health-practitioners, nurses, and doctors with an insightful view about the current discussion under the topics of side-effects and therapeutic processes, giving them an opportunity to monitor and validate the exchange of information in OHCs. We proposed hybrid models that combine high-level, abstract features extracted from convolutional neural networks with lexicon-based features and features extracted from long short-term memory networks to capture the semantics of the data. We show that our models, with and without lexicon-based features, outperform strong baselines.

ACKNOWLEDGMENTS

First of all, I would like to thank my advisor, Dr. Cornelia Caragea. She provided me with a perfect balance of guidance and freedom during my research. Thank you so much Dr.Caragea for your support and encouragement.

I express my sincere gratitude for Dr. Bill Buckles for accepting to be my major professor and showing my how kind-hearted and compassionate he is in dealing with students' problems during their academic journey.

I deeply appreciate my parents: Sareh and Muhammadjan, my sisters: Shahnaz and Mahnaz, and my brothers: Mehrdad, Bahman, and Mehdi. It is the passion for exploration and adventure combined with determination and hard work that I learned from you. Those values are what led me through my Ph.D. and let me have fun in the process.

Finally and foremost, I wouldn't be where I am today without the fantastic support, encouragement, and love from my beautiful and thoughtful wife, Somayeh, and my incredibly smart daughter, Adrina. I want you to know that none of this would have been possible without your support, patience and inspiring advise.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

In this chapter, we discuss the background and motivation of our study in Online Health Communities.

1.1. Background

With the advent of the Internet and advances in computer technology, the way that people accomplish their tasks or respond to their requirements has been revolutionized these days. Recently, patients are taking to the Internet more and more to find information [15]. For example, in the past, patients refer to their health practitioners, nurses or doctors to acquire information about their disease. However, Pew Research suggests that people spend more than 20 hours online per week, and most of them (i.e., 72%) had the experience of using the Internet for the health-related issues. Almost 60% of all US adults looked for someone else's health-related experiences or information about specific symptoms and treatment [37]. Inspired by the high volume of Internet usage, recently many blogs, forums, and social networks emerged in the health-related domain to help patients in supporting each other emotionally and exchanging information regarding their personal experience. For example, Breastcancer.org has created many discussion boards for patients who are suffering from breast cancer and also for breast cancer survivors; in a larger spectrum, Cancer Survivors Network (www.csn.cancer.org) provided an environment for many types of cancers from breast cancer to prostate and lung cancer, which include a large spectrum of cancer.

There are some key terms which are repeatedly used in the context of the health-related networks. To be able to understand established studies in this domain, we need to define some terms which repeatedly appear in the research articles.

(1) Social integration: This term refers to having social ties.

(2) Social network: An integration of social relationships that circle individuals. These networks represent any connections between people with any types of interactions between them (e.g., may or might not being supportive of each other) [44]. The social network creates social support.

(3) Social support: According to the seminal work by House [1981], social support is the functional content of relationships that can be categorized into four broad supportive behaviors (Figure 1.1) [5, 19, 41]:

- Emotional support which contains empathy, love, trust, and caring.
- Informational support which involves the provision of suggestions, advice, and information that a person can use to address challenges.
- Instrumental support which includes physical help that directly assists the person in need.
- Appraisal support which is the provision of information that is required for self-evaluation purposes.

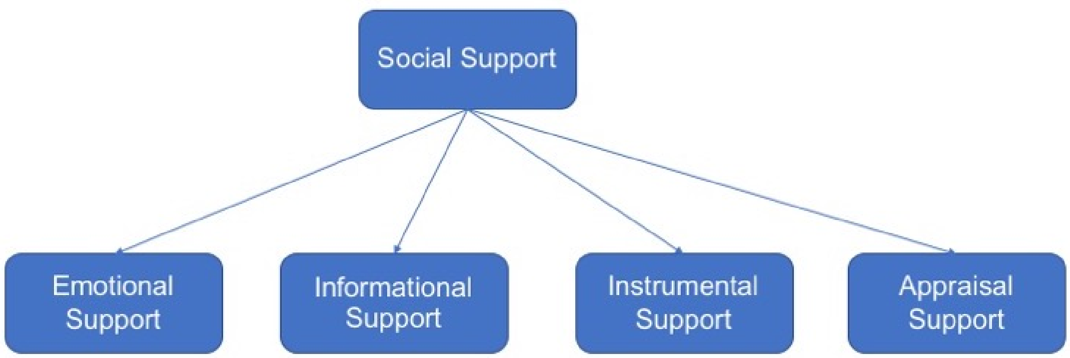

FIGURE 1.1. Types of social supports in OHCs' messages.

Social support providers deliberately provide a service to their audiences to support them in their lives. In other words, social support is a conscious behavior which is different from social influence exercised via observing others' behavior. OHCs provide an environment for their participants, and their family members and friends to share their experiences and information about prescribed medicines, side effects, therapeutic processes, mental health, and feelings. OHCs are deemed as having two principal functions: emotional support and informational support [23, 34]. Prior studies investigated the value of social support (i.e., emotional and informational support) and found that it brings about better feelings and fewer mortality odds to patients [49, 61].

1.2. The Importance of Research in Online Health Communities (OHCs)

As a result of patients' high contribution to the health-related online applications, an enormous amount of data has been created that contains details about users' opinions on different topics, individuals' interactions, and the content of their conversations. This data provides researchers with an unprecedented opportunity to study the content of online networks at a scale and granularity that was not possible before. Messages in Online Health Communities (OHCs) have been investigated for different purposes including the mechanism of social support creation [96], analyzing patients' behavior [71], or recognizing life-threatening disease [107]. Investigating OHCs' data also suggested that interacting with a larger group of people in OHCs is associated with lower mortality odds in people who are suffering from severe disease [48, 61]. Many studies postulated that the main reason for patients using OHCs is to receive emotional and informational support from their peers [25, 69, 123]. Studies also suggest that patients found OHCs a better resource for obtaining information in comparison with health professionals [43] and care providers [94]. The results of the research confirm that patients provide a better understanding of each other and provide practical instructions to their peers on how to manage any problem regarding their disease [7]. As can be seen, studies on messages extracted from OHCs has the potential to provide valuable information about patients' characteristics and many details about their behaviors.

1.3. Computational Models for OHCs

In the medical and social sciences, studies are based on interviewing patients, filling out questionnaires, and monitoring and conducting some controlled experiments on a couple of hundreds (at most) of samples [38, 45, 93]. These studies suffer from scalability, biased data usage [89], and high dependency on human memory that might not recall details accurately [66, 91]. However, in OHCs, thousands of people are contributing to share their valuable, first-hand information that contains therapeutic processes, types of medications and their effectiveness, side-effects and reactions to different drugs, emotional experiences from the time a patient learns about her/his symptoms, the story of accomplishments, and possible failures. Extracting information from user-generated data, which is created through everyday interactions between people, is shown to yield a more accurate understanding of their opinions in comparison with the self-report questionnaires

[13, 119]. In fact, analyzing user-generated comments, which is data-driven and can be updated over time, gives a more natural and accurate understanding of patients' opinion of health-related issues rather than providing patients with some pre-planned, structured questions that should be responded in a given period of time.

Many studies applied variety of machine learning, data mining, and natural language approaches for massive amount of data captured from general social media and Internet web page content [68, 90]. Recently, some studies [10, 89, 113, 126] developed computational models for analyzing OHC messages to analyze patients' behaviors and characteristics by using a very large amount of data which is unimaginable for non-computational studies.

1.4. Motivation and Contributions

Despite the fact that many studies have endeavored analyzing data at a large scale in OHCs, there is still a very big gap between computational studies' results and the real need for patients, OHC moderators and medical domain's experts. For example, prior studies provide a very high-level and general analysis of emotional support such that they only investigate the existence of emotion in a given OHC message [10, 113] without addressing the detailed information in emotional or informational messages; Informational messages contain valuable information regarding side-effects, therapeutic process, drugs, etc. that cannot be captured by a general analysis of messages in OHCs.

The main purpose of patients in using OHCs is to receive emotional and informational support [25, 69, 123], and the main purpose of developing OHCs [12] is:

- To provide high-quality and valid information efficiently to the participants.
- To understand the mechanism and the impacts of social supports (e.g., how social support forms, who has the most positive impact on other members? and why?, etc.) on patients.

However, current features and facilities in OHCs are quite inconsistent with the set goals and their structures are very primitive such that OHC participants (i.e., patients, medical researchers, practitioners, doctors, etc.) can only manually browse a variety of forums and search for their favorite information based on the topic of threads or using keyword-based search engines without the ability to mine data or provide abstract level information to its user.

The main purpose of this dissertation is to develop several computational models to analyze emotional and informational messages in OHCs in detail to help patients, OHCs' moderators, and researchers in the medical and psychological domain to access their favorite information from massive amount of data in OHCs efficiently. For example, we develop a computational model which is able to identify emotional messages accurately and furthermore detect the type of emotion in them. This model can be used by OHCs' moderators and psychologists for identifying influential users who spread positive feelings in OHCs and can also help depressed patients who post messages with sad content. In this dissertation, we address the following problems:

(1) Emotional support analysis
(2) Empathetic messages identification and analysis
(3) Informational messages identification and analysis

1.4.1. Emotional Support Analysis

We propose a model for detecting emotional messages in OHCs. Our computational model combines the strengths of convolutional and long short-term memory networks to uncover the semantics latent in text. Our contributions are as follows:

(1) We propose to detect emotion types in messages posted in online health communities. Identifying emotion types in patients' messages augments the capability of OHCs' moderators, caregivers, and doctors to provide high-quality services to OHCs' users or patients. To our knowledge, we are the first to address emotion type detection in OHCs.
(2) We propose a computational model, called ConvLexLSTM, for emotion detection in OHCs. Our model combines the output of a Convolutional Neural Network (CNN) with lexicon-based features, which are all fed into a Long Short-Term Memory (LSTM) network that produces the final output via softmax. We show empirically that ConvLexLSTM significantly outperforms strong baselines and prior works. Moreover, we show that the proposed model continues to perform well even in the absence of lexicon features.
(3) Finally, we applied ConvLexLSTM in a large scale experiment to study the correlation be-

tween US holidays and users' emotional states, which can help design smarter approaches to improve patients' moods.

(4) As part of our contributions, we constructed two datasets for emotion type detection in an OHC. To our knowledge, our datasets are the first constructed for this task.

1.4.2. Empathetic Messages Identification and Analysis

Empathy captures one's ability to correlate with and understand others' emotional states and experiences. Messages with empathetic content are considered as one of the main advantages for joining online health communities due to their potential to improve people's moods. Our contributions are as follows:

(1) We propose a machine learning model for identifying empathetic messages in OHCs. To our knowledge, this is the first work on automatically detecting empathy in OHCs.

(2) We experimentally validate our empathy identification model on a manually annotated dataset generated from the Cancer Survivors' Network of the American Cancer Society.

(3) We show that, in general, empathetic messages are correlated with a positive change in participants' sentiments.

1.4.3. Informational Messages Identification and Analysis

We propose to analyze messages in OHCs to extract the information type that they contain, i.e., *therapeutic procedures* and *side effects*. We design a computational model that is able to exploit the semantic information from text, and coherently combines high-level (abstract) features with surface-level and lexicon-based features.

(1) We propose to extract fine-grained information types from messages posted in online health communities. Identifying information types provides doctors, health practitioners, and OHCs' moderators with an insightful view of patients' physical status during various treatments. In addition, it can provide newly diagnosed patients with information about what they should expect throughout their treatments and help them in making informed decisions about their disease more effectively [98, 99]. To our knowledge, we are the first to address fine-grained information type extraction in OHCs.

(2) We design and explore a computational model that can identify messages belonging to *therapeutic procedures* and *side effects* with high accuracy. Our model is a hybrid neural network combined with lexicon-based features, which we call HNNL. HNNL combines the output of a Convolutional Neural Network (CNN) with the output of a Long Short-Term Memory (LSTM) network and lexicon-based features, which are all fed into a fully connected network with a SoftMax layer.

(3) We empirically show that HNNL significantly outperforms strong baselines and prior works; moreover, we show that the proposed model continues to perform well even in the absence of lexicon-based features.

1.5. Dissertation Outline

In what follows, we provide a brief description of the chapters in the dissertation. Each chapter corresponds to a paper. Some of the research work in the dissertation has been published in conference proceedings, is under review in a conference proceeding, or is in the preparation process. The aim of this research is to propose some computational models to automatically classify and extract OHCs' messages based on the concept latent in them. These messages can be used by patients (e.g., for identifying side-effects of a specific drug), OHCs's moderators (e.g., for identifying patients who spread out joy emotion in OHCs), psychologists (i.e., to identify depressed people based on the sad messages that is posted by patients), or doctors (i.e., to validate information regarding side-effects and therapeutic process which are provided by patients in OHCs), medical domain researchers (e.g., to investigate relations between posting empathetic messages and the patients characteristics like being an influential user), and others.

Chapter 2 presents our proposed computational model for analyzing emotional support in OHCs. We extend the dataset which was provided by previous studies in three ways: First, in terms of the number of annotated sentences in the dataset; in the previous study there were 1066 annotated sentences; we annotate 1047 new sentences and our dataset has 2107 sentences in total. Second, all sentences in the previous dataset were harvested from a single forum, i.e., a breast cancer forum; we capture all newly-added-sentences from the lung cancer forum which provides the potential for new studies like studying differentiation between gender behaviors (the distribution of genders are very different in the breast cancer and the lung cancer forum). Third, prior studies annotated sentences for emotional vs. non-emotional categories; we extend the domain of the annotation to detect the type of emotion (i.e., joy, sad, surprise, fear, angry, and disgust) in each sentence if they are emotional. As a proof of usefulness of annotating sentences from different forums and also the usefulness of annotating emotion type in messages, we show that men are more emotional and send more messages with joy content when they post a message in a forum in which the majority of its members are female (i.e., breast cancer, where the majority of its members are female in comparison with the lung cancer forum). We then improve the previous computational model in identifying emotional messages in OHCs. We develop a computational model which does not need to use any handcrafted feature. We are the first study that proposes a deep neural network for developing a computational model in analyzing messages in OHCs. We show that our proposed model outperforms traditional models even without using any handcrafted feature; this implies that our model can be applied to messages from any forum in OHCs. In OHCs, the content of messages are quite domain dependent with medical specific expression which has a domain-specific meaning(s) [33]. Finally, by using our proposed models, we investigate the emotional state of patients in an OHC (i.e., Cancer Survivors Network) around important holidays in the US.

Chapter 3 presents an approach for detecting messages with an empathetic content. As the first study on analyzing OHCs' messages for detecting empathetic messages, we create a dataset with 2107 messages which were captured from two different forums (i.e., breast cancer and lung cancer forums). We create word embedding vectors by using the whole messages that were posted over 10-years on the Cancer Survivors Network. We design and develop a computational model which is a deep neural network with a compositional structure. We apply CNNs network to capture semantic information latent in messages and used those feature to be injected into an LSTM network. We also generate several features including bag-of-words, tf-idf by using part-of-speech tags, and lexicon-based features which were used in an SVM classifier as baselines. We experimentally show that our proposed model for identifying empathetic messages is highly qualified. By using our model, we investigate the impact of empathetic messages on OHC participants. We experimentally show that empathetic messages in OHC messages transform participants' feelings from negative to positive.

Chapter 4 presents a computational model which classifies OHCs' messages as stating a side-effect, explaining a therapeutic process, or others. We annotate 2107 messages which were extracted from an OHC (i.e., Cancer Survivors Network) for this task. Our research framework in this chapter is the first study of its kind in the literature. In this chapter, we present a compositional neural network that extracts abstract and high-level features and combines them with lexicon-based features and the sequential features which are captured from LSTM network. We show that our model, with or without lexicon-based features, outperforms strong baselines. By using our models on OHCs messages, patients learn about their peers' experiences on side-effects efficiently, and they foresee and prepare themselves for those situations. In addition, our model provides patients with therapeutic processes which have been discussed already in OHC forums; by reading a collection of related messages, patients are not required to browse manually and read through many unrelated posts which is not as effective as what a computational model with high accuracy can do automatically.

Chapter 5 summarizes and concludes the dissertation and provides a summary of contributions and directions for future research.

1.5.1. Published Work

- *Analyzing social support in online health communities* has been published in *AAAI 2017* [56] with experiments on emotional-messages identification in OHCs. We performed a comparison with the state-of-the-art model and showed that our model performed considerably better than the prior model. Also, we have extended the emotional messages identification in some ways: by annotating new sets of sentences that was extracted from lung cancer, by proposing a new computational model for detecting the type of emotions latent in emotional messages in OHC, and by running several experiments to show the benefit if using emotion type detection in OHCs. We applied our emotion-type-detection model to the whole data in CSN to identify patients' emotional state around important holidays in the US. We also applied our computational models to the datasets that contain influential users' posts in CSN's breast cancer forums to analyze influential users' messages in terms of emotion types that is expressed in their posts. The extended version of the work is submitted to *ACL 2018* and is under review.

- *Identifying Empathetic Messages in Online Health Communities* has been published in conference proceedings of *International Joint Conference on Natural Language Processing (IJCNLP), 2017* [55]. This work is the first computational study on empathetic messages in OHCs. In this work, we drive a computational study on messages in OHCs. We show that our computational model, which captures semantic features from the text, outperforms strong baselines. We also show that empathetic messages improves patients emotional health and transform their sentiments positively (these results are all based on the data-driven approach). A journal version of the work in this chapter is under preparation, and we extend the work by investigating the relation between patients' emotional dynamics with the number of empathetic messages they read, showing patients' emotional dynamics when receiving empathetic messages from influential users and investigating its difference with receiving messages from regular members of OHCs. The outcome of this study is to help OHCs moderators and psychologists in detecting people who are spreading empathetic messages and assign them to people who has pessimistic view on their future. This study also has provided required ingredients for researcher who are working on influential users identification.

- *Analyzing informational messages in OHCs* has been submitted to conference proceedings of *SIGIR, 2018* and is under review. This work presents, for the first time in literature, a computational study for identifying messages explaining side-effects and therapeutic procedures. We propose a model that mines patients' messages in OHCs to identify positive experiences and suggestions described by other patients. This model combines high-level, semantic features with lexicon-based features and features extracted by the LSTM network. We show that our model outperforms strong baselines. We are currently working to extend this work by comparing the content of informational messages posted by influential users to the content of informational messages which are posted by regular users. The outcome of this research is to help patients efficiently retrieve and read through messages that explain side-effects or therapeutic processes.

1.5.2. Other Published Work (Not Included in This Dissertation)

- *Dialogue Act Classification in Domain-Independent Conversations Using a Deep Recurrent Neural Network*, in COLING (2016) (joint work with Nishitha Guntakandla, and Rodney Nielsen) [57]. In conversational text analysis, Dialogue Act (DA) identification plays a key role that shows the intention of the speaker in generating (speaking) each utterance (originally, utterance is a sentence); in other words, using a specific sequence of words when talking. The importance of DA identification increases, especially, when the conversation occurs is an open-domain, everyday conversation and not a task-based one. Initially a list of more than 200 DAs were generated in some literatures that was categorized into 42 and 5 DAs in the SwitchBoard and MRDA datasets. Some example of DAs are opinionated utterance, statement, WH-questions, etc. In this work, we present a deep recurrent neural network for classifying DAs. We showed that using Dropout method for regularization in LSTM networks, in the similar way that it is used in CNNs, decreases the performance of the LSTM networks. This work sets a new benchmark in classifying DAs in open-domain conversations and outperforms all prior traditional and DNNs-based models. This model is currently the state-of-the-art model in the domain.
- *Towards a Top-down Policy Engineering Framework for Attribute-based Access Control*, in SACMAT (2017), (joint work with Masoud Narouei, Hassan Takabi, Natalie Parde, and Rodney Nielsen) [84].

  Attribute-based access control (ABAC) is a kind of logical access control approach where authorize users of a network in doing a specific set of operations based on their attributes. These operations can be a permission on accessing a set of objects , the environment, and a number of sources that may be related to a current request. The benefit of developing a reliable ABAC to the enterprise space is that it promotes information sharing throughout an enterprise network while ABAC manages accessing to the network information. Many companies have high-level requirement specifications that specifies security policies including access control policies. IN this work, we take advantage of access control policies and present a top-down policy engineering framework for ABAC. The main goal of our work is to automatically extract policies from unrestricted documents that contain access policies. We then present a deep neural network classifier to extract sentences which contain a description of a access policy. IN this work, we build a dataset by annotating 2660 sentences from multiple organization policy documents. We applied our proposed model and the state-of-the-art model to compare their performances. We experimentally showed that our model outperforms the state-of-the-art model on our constructed datast as well as other datasets in the domain.

- *Identification of Access Control Policy Sentences from Natural Language Policy Documents* in IFIP Annual Conference on Data and Applications Security and Privacy, (2017), (joint work with Masoud Narouei and Hassan Takabi) [83].

  For any application which is running on a network, access control configuration are a necessary and essential element. Recently, many plausible access control models have been developed in information security domain. However, a variation of an standard access control, attribute-based access control (ABAC), has emerged that eliminated the limitations of the prior access control models while preserving their advantages. However, developing the attributed-based access controls is very dependent to information (e.g., policies) which is latent in a large amount of documents; Hence, a powerful natural language tool (e.g., information extraction tool) is required. Usually, organizations' policies are latent in the software requirements and policy documents which have a very big size and are mixed with many unrelated, general descriptions. Hence, manual processing of these documents is impossible or very expensive and challenging. In this work, we present a computational model for identifying sentences with ABAC content. In our computational model, we used several natural language processing tools and techniques such as pointwise mutual information metric, part-of-speech tags, and syntactic relations between words. We show that our computational model is domain independent; we evaluate our model on policy documents collected from different domain including healthcare, conference management, and education. Our experiments show that our model outperforms state-of-the-art models.

CHAPTER 2

EMOTION DETECTION IN HEALTH-RELATED ONLINE POSTS

Detecting and analyzing emotional messages in Online Health Communities (OHCs) can provide insightful information about patients' emotional states. Current approaches to emotion analysis in OHCs focus only on identifying messages that contain emotions with no emphasis on emotion type detection such as joy or sadness. In this chapter, we show how high-level and abstract features derived from deep learning models can be employed for improving prior results and strong baselines on emotional message identification as well as emotion type detection, e.g., *joy* and *sadness*. Specifically, we propose computational models for emotion detection that combines the strengths of convolutional and long short-term memory networks to uncover the semantics latent in text, and show that our models, with and without lexicon-based features, outperform strong baselines.

2.1. Introduction

Emotions are an essential part of our lives and reflect feelings such as joy, sadness, and fear, which affect our overall well-being. Emotion detection from text has been extensively studied from news headlines, social media, and song lyrics using computational models [54, 80].

Recently, emotion detection has started to emerge in online health communities (OHCs) [10, 114]. OHCs provide a user-friendly environment for patients, families and friends to share thoughts and socialize with each other on various topics such as therapeutic processes, prescribed medicines, side effects, and mental and emotional health. Emotional communications in OHCs is considered as one of the main advantages of using OHCs that causes patients to feel better [124, 125] and decreases the odds of mortality in patients [61]. Research has shown that patients who suffer from life-threatening diseases (e.g., cancer or AIDS) feel emotionally supported when they interact with their peers in OHCs [89, 124, 125]. Table 2.1 shows an example of communication between two patients in an OHC where one patient emotionally supports her peer and as a result, the initiator of the thread feels better. Emotional support comprises of seeking or providing caring/concern, understanding, empathy, sympathy, encouragement, affirmation, and validation.

–***Patient:*** *Hi all sense being on chemo (5 down 1 to tch) with the last two really I have had a problem with my BP being high. I am having a problem with my heart racing. At rest it may get down to 86. When my* oncologist did the muga scan it *went from 68 to 63. I have never had a problem with my heart at all. I'm Very nervous.*
–***Commentator***: *I had much the same problem while doing chemo, the last 2 or 3 rounds were the worst. not to worry to much! By the way I am the proud owner of 3 chihuahuas. Blessings to you...Alison*
–***Patient:*** *Thanks so much I feel allot better now. I did talk to my Dr and he is giving me meds to lower the rate. I feel like I spend my time fighting side affects LOL. Thanks sisters. Take care all*

TABLE 2.1. A sample of an emotional comment.

"My doctor's office is very clean, who cares when he has prescribed me a wrong medication for six months!"

TABLE 2.2. Example of an emotional message from an OHC that contains a sad emotion.

Despite the importance of emotion detection in OHCs, most of the recent works have been devoted to high level emotion analysis, i.e., identifying messages that contain emotions, with no emphasis on the unique challenges associated with emotion type detection, e.g., joy or sadness. A deep understanding of the text and the writer's intention is required in order to correctly detect the types of emotions that are presented in messages posted in OHCs. Table 2.2 shows an example of a message that contains a sad emotion, which is hidden in the text. Interestingly, we ran several sentiment analysis tools on the message in Table 2.1, including Stanford CoreNlp [100], and all showed a positive sentiment, while the emotion expressed is clearly one of sadness. Thus, even using a coarse granularity approach to predict the negative sentiment of the message would not suffice.

In this chapter, we propose to analyze messages in OHCs to detect emotion types, e.g., joy or sadness. To achieve this, we design a computational model that is able to exploit the semantic information from text, and coherently combines high-level (abstract) features with surface-level and lexicon-based features. Our contributions are as follows:

(1) We propose a computational model for identifying *emotional messages* in OHCs that eliminates the need for handcrafting features or building expensive lexicons that require experts' knowledge. Our proposed model called *deep convolutional LSTM*, or *ConvLSTM* for short, is described in the next section.

(2) We propose to detect emotion types in messages posted in online health communities. Identifying emotion types in patients' messages augments the capability of OHCs' moderators, caregivers, and doctors to provide high-quality services to OHCs' user or patients. To our knowledge, we are the first to address emotion type detection in OHCs.

(3) We propose a computational model, called ConvLexLSTM, for emotion type detection in OHCs. Our model combines the output of a Convolutional Neural Network (CNN) with lexicon-based features, which are all fed into a Long Short-Term Memory (LSTM) network that produces the final output via softmax. We show empirically that ConvLexLSTM significantly outperforms strong baselines and prior works. Moreover, we show that the proposed model continues to perform well even in the absence of lexicon features.

(4) Finally, we applied ConvLexLSTM in a large scale experiment to study the correlation between US holidays and users' emotional states, which can help design smarter approaches to improve patients' moods.

(5) As part of our contributions, we constructed two datasets for identifying emotional messages and emotion type detection in an OHC. To our knowledge, our datasets are the first constructed for this task.

2.2. Data Collection and Annotation

We used two datasets for evaluating our model. The first dataset is provided by Biyani et al. [2014], which contains 1066 messages from the breast cancer discussion board [1], denoted as B-DS. Biyani et al. [2014] considered each message as a single sentence of each comment in OHCs. They used sentences of comments because comments are usually very long and users talk about different topics in each comment. However, on the sentence level analysis, we can make a better estimation on the purpose of the commentator in writing that sentence, whether or not

[1] www.csn.cancer.org

expressing his/her emotion. For building the second dataset, we randomly selected 225 comments from 21 discussion threads in the lung cancer discussion board in the Cancer Survivors' Network (CSN) of the American Cancer Society. We denote this second dataset as L-DS. Following Biyani et al. [2014], we extract all sentences out of comments and shuffle them and choose messages with a length greater than four words. We ended up with 1041 messages for L-DS. In total, we are provided with 2107 labeled sentences.

Lung cancer and breast cancer discussion boards have the two largest discussion boards among all CSN's sub-forums. CSN is a popular OHC which provides friendly environments for patients, survivors, and their family and friends.

The purpose of our annotation task was two-fold: firstly, to identify emotional messages in L-DS, which is similar to what was done in Biyani et al. [10]; secondly, to determine the emotion types (e.g. anger, disgust, fear, joy, sadness, and surprise) in messages. The annotation task was conducted in an iterative fashion following prior studies and guidelines [31, 35, 97]. We used the B-DS for the training purposes when annotating L-DS for emotional messages. In each round of the training phase, two annotators were given 100 instances of messages from B-DS, and each annotator was expected to achieve at least 70% pairwise agreement with the provided annotation to be able to annotate L-DS. At the end of each phase, annotators met in groups with the researchers and discussed disagreements. After three rounds of the training phase, our annotators achieved 85% agreement with the annotated messages provided in B-DS. Upon passing the training period, annotators were assigned to annotate all 1041 messages from L-DS, and they ended up with 83% agreement in a period of two weeks. For the remaining 17% of messages, the annotators expressed their views on each case in the presence of the researchers and finally 100% agreement was achieved. Following this annotation scheme, we ended up with 655 (62.9%) emotional and 386 (37.1%) non-emotional messages in the L-DS. This distribution is almost similar to the distributions reported in Biyani et al. [10] for the B-DS, i.e., 63.41% emotional and 36.79% non-emotional messages.

In the second round of annotation, we used both datasets (i.e., B-DS and L-DS) to annotate the type of emotions in each emotional message. For emotion type annotation, we followed the

Emotions	Lung	Breast	Percentage(%)
Anger	59	29	4.0
Disgust	4	1	0.2
Fear	33	39	3.4
Joy	368	470	39.7
Sadness	183	134	15.0
Surprise	8	3	0.5

TABLE 2.3. Emotion distributions in B-DS and L-DS with the percentage of each emotion in the aggregated version of L-DS and B-DS.

six emotions categories suggested by Ekman's [32] including anger, disgust, fear, joy, sadness, and surprise. Our annotators were allowed to attribute one or more emotions to each message. For instance, a message could be annotated as bearing *sadness* or a combination of *sadness and fear*. In each round, 300 messages were assigned, and we asked annotators to meet with the researchers in a group to discuss disagreements and document their discussion before the next 300 instances being assigned. We assigned the B-DS and the L-DS to the annotators gradually through several rounds and 100% agreement being achieved via several discussions during 20 days. Table 2.3 represents the distribution of emotions in 2107 instances. As we can see, joy and sadness have the greatest distribution in the whole dataset and in each dataset individually. This distribution makes sense due to the fact that people usually express their intense feelings like disgust with the ones with whom they have a close relationship. One of the reasons for high percentage of *sadness* and *joy* distributions are the number of *sympathetic* and *empathetic* messages in OHCs.

2.3. Model

In this section, we describe our models: ConvLSTM and ConvLexLSTM, which have the same structure with a small variation in ConvLexLstm that uses lexical features for the cases in which lexicon is provided.

2.3.1. Convolutional Neural Network

Convolutional Neural Network (CNN) is one of the main deep neural network architecture which is inspired by the natural perception structure of the human brain [51]. Figure 2.1 shows the structure of CNN networks that apply convolutional functions on the vectors of sentences. In general, convolution function is defined as follows:

$$s(t) = \int x(a)w(t-a)da \tag{1}$$

which is typically denoted as follows:

$$s(t) = (x * w)(t) \tag{2}$$

where the asterisk represents convolution function at time t. In the terminology of the convolutional neural network, the first parameter is **input** (i.e., x), and the next argument, w, is called a **kernel** or a **filter** interchangeably. Finally, the output of the equation is called a **feature map**. As we can see, a feature map is not dependent to the prior or future time step; Thus, features extracted by convolutional network are time invariant.

As can be seen from Figure 2.1, CNN networks consist of three general phases:

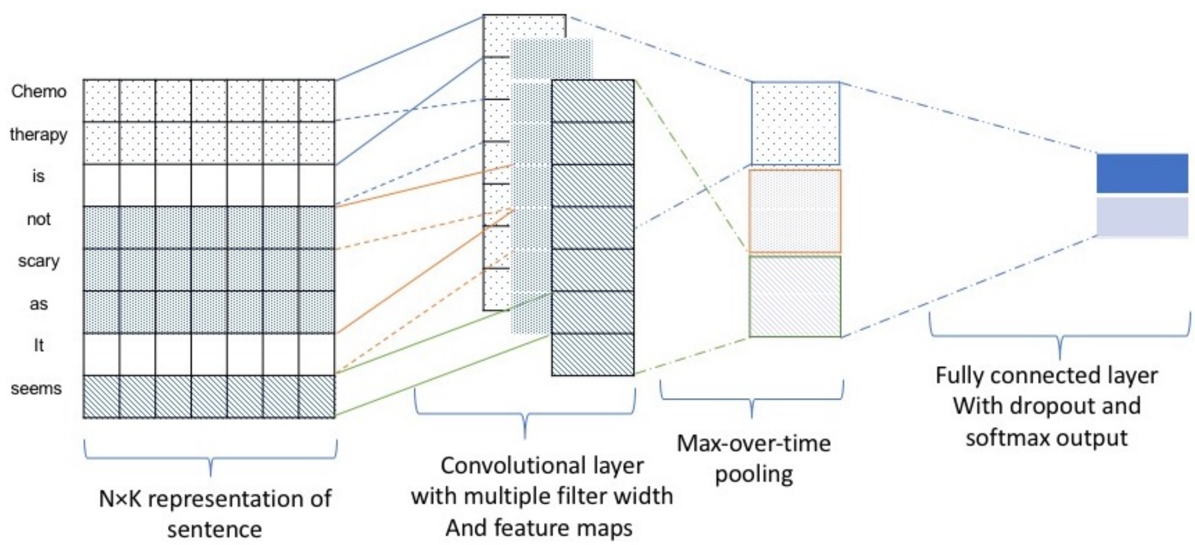

FIGURE 2.1. The structure of convolutional neural network (CNN) that applies multiple filters on a sentence with N tokens and K vector length.

- Through the first phase, multiple convolutions is performed concurrently. The convolutional functions variables are comprised of a set of learnable kernels that create some set of linear activation. Through the initial steps, each kernel's size is small and covers handful of words in a sentence (i.e., one, two, three, etc. words at each step). Kernels slide or convolve through the height of the input sentence and measure dot products between figures, pre-assigned randomly in the kernel and the input vector of each word in the sentence. Sliding kernels will generate a linear activation map
- Through the second phase, every linear activation is run by a nonlinear activation function like rectified linear function. Intuitively, kernels will learn to activate when they visit a combination of some bi-grams or tri-grams. These learned kernels are stacked along the depth dimension and create multiple sets of activation maps.
- Finally, the last phase is the pooling function that changes the resulting figures of the layers. In this phase, the feature map is created.

Pooling functions are a kind of statistical functions that yields a summary of statistics of neighboring outputs in a layer. For example, Max pooling function yields the maximum figure among all generated outputs of each kernel. *Average pooling* and *1-max pooling*, which is the maximum of the last state of kernels' output, are the other popular types of pooling in CNN networks.

As can be seen, Convolutional neural networks (CNNs) learn local features and assume that these features are not restricted by their absolute positions. Although CNNs do not preserve location information of each word in sentences, many studies in natural language processing (NLP) have used CNNs successfully in their studies for statistical language modeling [26, 59, 110], text classification [21, 22, 122], and Named Entity Recognition (NER) [17, 63]. We also use CNNs in combination with Recurrent Neural Network (RNN). We employ CNNs as an automatic feature extractor which is able to create high-level, abstract features that address semantic part of the text.

2.3.2. Recurrent Neural Network (RNN)

RNNs [95] are a type of neural network which is developed for processing sequential data which have the values of $X_1, ..., X_t$. RNNs have a very similar network architecture to the Feedforward neural network, i.e., Figure 2.2, except that each connection has an extra backward connection Figure 2.3.

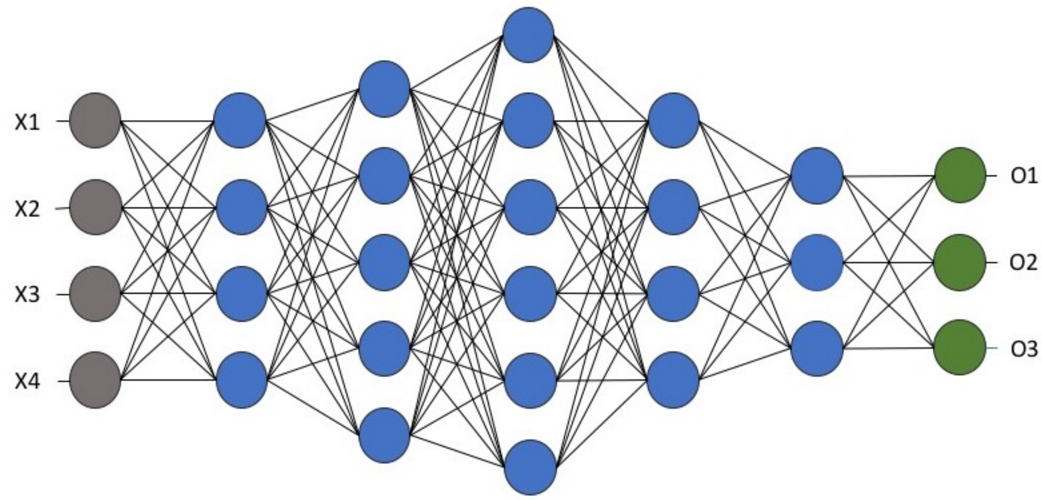

FIGURE 2.2. Feed-forward fully connected neural network

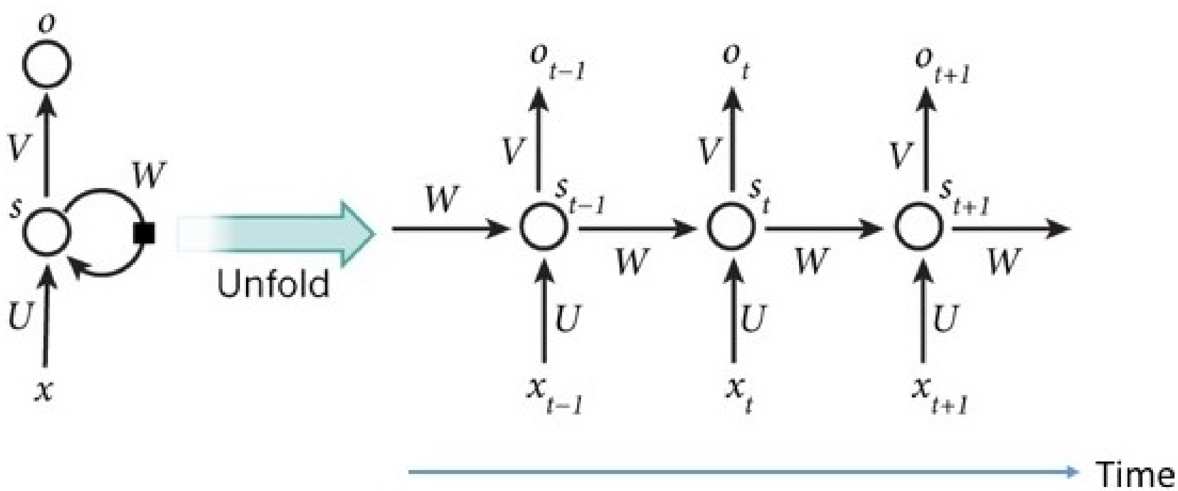

FIGURE 2.3. One-layer recurrent neural network

As can be seen from Figure 2.2, at each *time step* (sometimes called *frame*) a neuron receives two inputs: x_t and y_{t-1}. y_{t-1} is the x_t's output from the previous time step. Figure 2.2 represents RNNs against the time axis.

For a single instance the output of a RNN layer is:

$$y_t = \phi(W_x^T \cdot X_t + W_y^T . Y_{t-1} + b) \tag{3}$$

where W_x and W_Y are weights for X_t and the output of the previous time step (i.e., Y_t), respectively. ϕ represents activation function (e.g., Relu, Tanh, etc.).

Theoretically, RNNs preserve the record of previous incidents, but in practice, when the gap between related information extends, RNNs fail to maintain relevant information. Hochreiter [1991] and Bengio et al. [1994] investigated the main reasons for RNNs' failures in detail. The two major problems with RNN is the vanishing and exploding gradient that causes the learning process to be terminated prematurely [76, 88]. These two problems appear when gradient descents are applied subsequently in many time steps that causes RNNs to lose information of previous events; this effect contradicts the main purpose of using RNNs. For example, consider the following message: *"I wasn't able to joke with the nurses and doctors through all my treatments, but once I was done and the tamoxifen was in my system, I could even laugh ."* If we apply RNNs for the classification purposes for this example, we need 30 time steps; this means that as a result of the vanishing/exploding effect, RNNs give higher weights to the last part of the text (i.e., I could even laugh) while it forgets signals coming from the first chunk of the text (i.e., I wasn't able to joke).

2.3.3. Long Short Term Memory Network (LSTM)

LSTM networks are a variation of RNNs that are developed to solve RNNs problems [47]. We use LSTM networks in developing our models, which are tuned to capture long-distance dependencies as their default specificity. In classifying of messages in OHCs, having the ability to connect related expressions of information that are distant from each other is important, particularly when it comes to classifying messages as either subjective or objective (emotional messages are subjective versus informational messages which are objective). Classifying subjective versus objective texts is one of the major tasks in sentiment analysis in which LSTM-based approaches are shown to achieve high-quality results [100]. Another reason for using LSTM networks is that

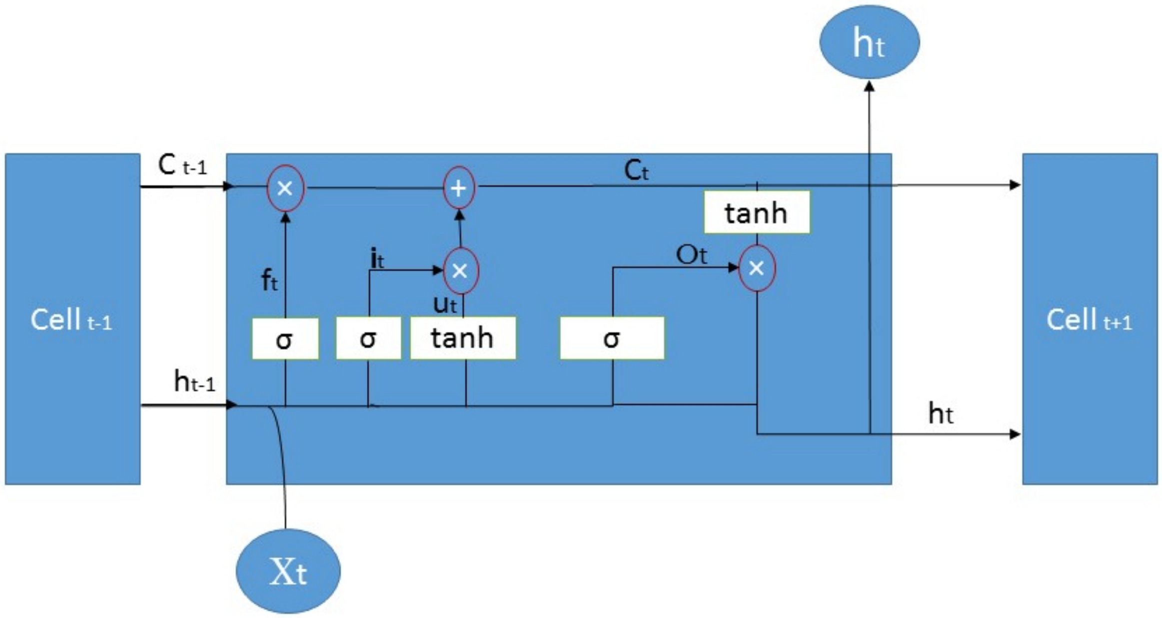

FIGURE 2.4. LSTM cell structure and its respective parameters

they use a *forget gate layer* (see Figure 2.4) to distill trivial weights, which belong to unimportant words from the previous cell states (see Eq. 4) .

Figure 2.4 illustrates a standard structure of an LSTM cell. LSTM networks have been shown to improve state-of-the-art models' results in many text classification tasks [3, 52, 116]. In our task, we also need to develop a model which can preserve distant information in OHCs messages, especially when it comes to emotion detection tasks where a person might describe his/her emotional state and later in the message he/she negates that emotional state (see the previous example). LSTM networks can learn what information should be preserved and what information should be thrown away. The long-term signals are carried out by $c_{(t)}$ and the short-term signals are carried out by $h_{(t)}$. An LSTM cell has three gate controllers that amplify or weaken $c_{(t)}$ signals:

- The input gate (Eq. 4) manages how much of $u_{(t)}$ should be added to $c_{(t)}$.
- The forget gate (Eq. 5) manages the amount that should be eliminated from $c_{(t)}$.
- The output gate (Eq. 6) decides which part of $c_{(t)}$ should be added to the current state and $y_{(t)}$

An LSTM can be considered as a type of deep recurrent neural network (RNN) with an indefinite number of layers that preserves information from the previous time steps (e.g., words sequences, words locations, etc.). The following shows equations for computing and managing signals:

(4) $$i_{(t)} = \sigma\left(W_{xi}^{T} \cdot x_{(t)} + W_{hi}^{T} \cdot h_{(t-1)} + b_i\right)$$

(5) $$f_{(t)} = \sigma\left(W_{xf}^{T} \cdot x_{(t)} + W_{hf}^{T} \cdot h_{(t-1)} + b_f\right)$$

(6) $$o_{(t)} = \sigma\left(W_{xo}^{T} \cdot x_{(t)} + W_{ho}^{T} \cdot h_{(t-1)} + b_o\right)$$

LSTM unit at time t computes the memory cell:

(7) $$u_{(t)} = \tanh\left(W_{xu}^{T} \cdot x_{(t)} + W_{hu}^{T} \cdot h_{(t-1)} + b_g\right)$$

(8) $$c_{(t)} =_{(t)} \bigodot u_{(t)} + f_{(t)} \bigodot c_{(t-1)}$$

and then computes the output, or activation:

(9) $$y_{(t)} = h_{(t)} =_{(t)} \bigodot \tanh(c_{(t)})$$

where:

- W_{xi}, W_{xf}, W_{xo}, and W_{xg} are the weights of four controllers when connecting to input vector (i.e., x˙(t))
- W_{hi}, W_{hf}, W_{ho}, and W_{hg} are weights of four controllers when connecting to the previous hidden state (i.e., h˙(t)).
- b_i, b_f, b_o, and b_g are the bias parameter for each of four controllers.

The resulting sequence of the cells is $h_1, h_2, \cdots, h_T$, where T represents the length of the input sequence.

In our work, since we need an abstract and high-level features of each message in OHCs, we use LSTM networks which have multiple layers. Figure 2.5 shows the structure of LSTM cells in multiple layers. From the bottom layer to the top, multi layer LSTM networks generate more abstract and high-level features. We apply the softmax function at the very end layer for classification.

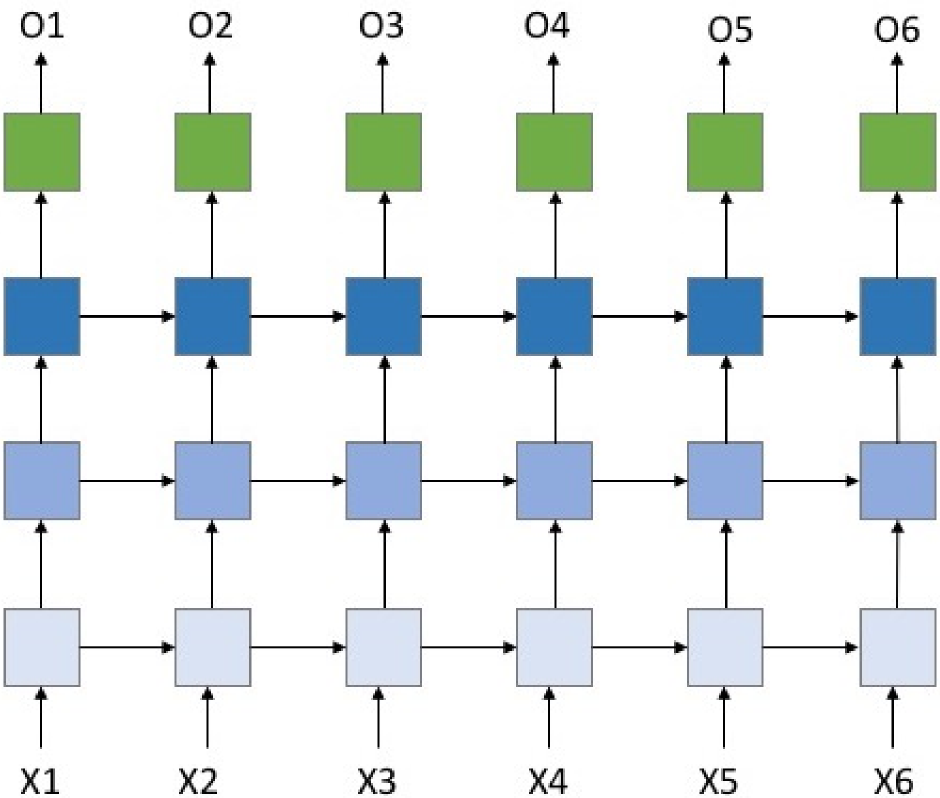

FIGURE 2.5. Multi-layer LSTM networks structure

Recently, it is common to use a more complicated variation of the RNN model, such as an attention and a memory based model, TDLSTM+ATT [2]. This is an extension of LSTM networks that use external memory and attention mechanisms for processing long texts, with a high time-complexity [106].In addition, TDLSTM+ATT has more hyper-parameters and needs a bigger training set (which makes them inappropriate for small or mid-size datasets) compared with

LSTM networks.

2.3.4. Word Embedding

To be able to work on text data, we need to choose a word representation. One-hot vector is an option for an initial setting. However, One-hot vector provides a very sparse representation that will not be efficient. For example, if the size of our vocabulary is 100,000 words then the n^{th} word will be represented as a vector of the length as 100,000, where only the n^{th} element is one, and all other elements equal zero. Using embedding vectors is a solution to avoid inefficiencies introduced by One-hot vector. Instead of using One-hot vector, we can assign a dense vector with a reasonable length (i.e., 300) to each word in the vocabulary and let the deep neural network gradually learn the best embedding for each word via a training phase. Creating the initial embedding vectors for each word in the vocabulary can be done by using one of the following methods:

- Initializing a vector for each word in the vocabulary and let the neural network to learn the best figures during training phase.
- Using one of the methods (e.g., Word2vec) used for distributed representation of words.

Many studies postulated that using Word2vec [77] results in a better result in comparison with random initialization of embedding vectors [58, 126]. In our work, we used Word2vec for initializing embedding vectors that are used as input vectors in our networks. There are two main parameters for creating word embedding vectors: dimension, and resources. We used Wikipedia [2] and CSN (the whole CSN data from June 2000 to June 2012 with volume size of 752.4 MB) as our resources. Wikipedia is a rich resource with medical information including cancer, therapies, medications, etc. We generate word embedding vectors by applying the W2vector module in *Gensim* [92]. Table 2.4 shows the top-10 most similar terms to the word ”cancer,” using Wikipedia (wiki-150) and CSN (csn-150) as resources with 150 vector dimensions. The Top-10 most similar words also illustrate the ranking of the most similar terms to the word ”cancer” from the top left to the bottom right. In the ranking representation, we ignored any variations of the word ”cancer” (e.g., cancers, canzer, cancer, etc.). As we can see, two out of the three most similar terms in

[2] https://dumps.wikimedia.org/enwiki/20160421

both word embeddings are the same, *melanoma* and *lymphoma*. However, from the third term in wiki-150, the similar terms gradually skew towards the general semantic of the word "cancer," that represents cancer as a type of disease. However, in the csn-150, we can observe that all terms are the name of the different types of cancer including abbreviation e.g., bc (Breast cancer) or ovca (Ovarian cancer). These observations suggest that the word embedding created by Wikipedia is prone to present a general view of the health-related data rather than the one generated by CSN data. We apply both word embeddings in experiments to find out which resource turns out to be the best for the classification task.

CSN-150	melanoma, lymphoma, bc (short form of *Breast cancer*), ovca (short form of *Ovarian cancer*), disease, adenocarcinoma, tumour, leukemia, metastasis, carcinoma, dysplasia, malignancies, NHL, adenocarcinoma, UPSC
Wikipedia-150	leukemia, melanoma, lymphoma, diabetes, tumor, tuberculosis, disease, pneumonia, polio, meningitis, cirrhosis, hepatitis, mesothelioma, emphysema, alzheimer

TABLE 2.4. The most similar terms to the word "cancer"

2.3.5. Lexicon-Based Features

The two principle functions of online health communities are emotional and informational support. Accordingly, we expect to see more of messages in OHCs which convey one's caring, feelings, sympathy, etc. to other members or to see information about medications, disease report, side-effects, referrals etc.. As we can see, emotional and informational messages express different types of concepts that need the commentator to use different sets of words in each of them. For example, it is more likely to find subjective words in emotional messages rather than in informational messages. Informational messages use keywords that are more related to a disease.

Lexicon-based approaches for detecting emotions in the text have been the most widely-used features of many models [67, 79, 102, 104]. As we can see from Figure 2.6, there is a pretty good chance that a commentator uses an emotional keyword when he/she provides support to

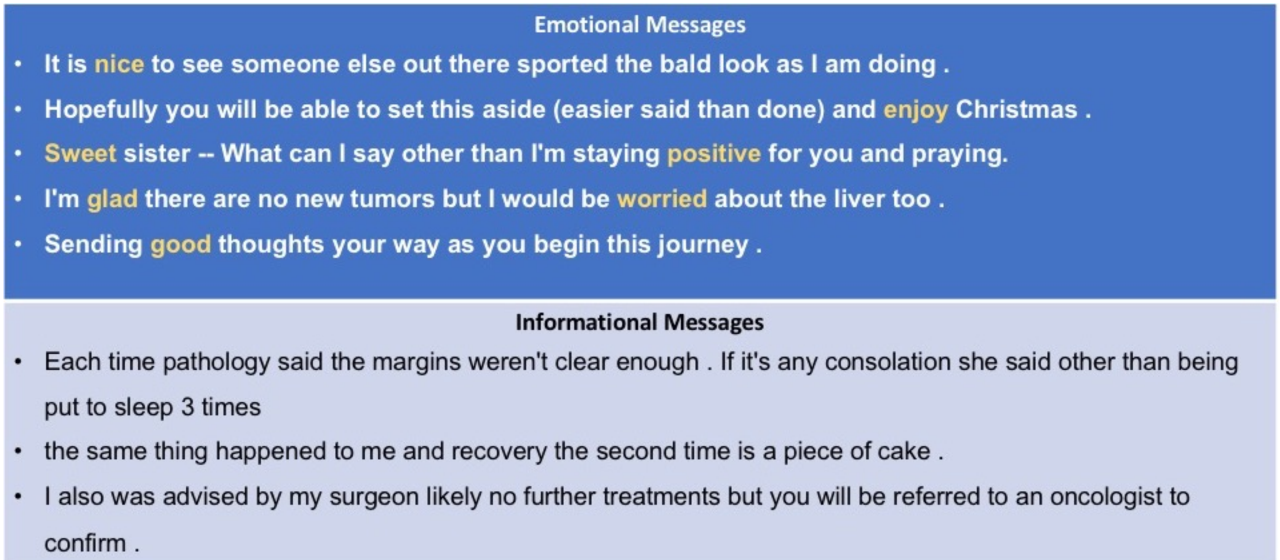

FIGURE 2.6. Difference between emotional and informational supports in terms of content words

other members of an OHC. We created seven lexicons to address differentiation in the content of emotional and informational messages; these lexicons are denoted as: **EmoLex1**, **EmoLex2**

- **numWeak**: This lexicon contains weakly subjective words which are words with specific subjective usages. These words were extracted from the MPQA subjectivity dataset [115] and a similar lexicon that was provided by Biyani et al. [2014].
- **numStrong**: This lexicon contains strongly subjective words that are subjective in most contexts. This lexicon is compiled from the MPQA dataset [115] and a similar lexicon provided by Biyani et al. [2014].
- **numDrug**: This lexicon contains medications and drug names that are more related to breast and lung cancer. This dataset is an extension of the lexicon that was created by Biyani et al. [2014]. We extend this lexicon by gathering information from online resources [3]. Our assumption is that drug names appear mostly in informational messages.
- **numSide**: Many patients in OHCs share their experiences on taking drugs and their respective side-effects. Many therapeutic processes also have major side-effects especially in cancer treatment processes (e.g., chemotherapy). We used online sources to capture

[3] An example of online source: http://www.breastcancer.org/treatment/druglist

common side-effects in cancer treatments [4]

- **numProc**: We used name of therapeutic procedures that patients commonly use when they exchange information with each other. We extend a lexicon created by Biyani et al. [2014] using online resources.

The above-listed lexicons address differentiation between emotional and non-emotional messages. However, we need more granular information for differentiating between a variety of emotion types. Hence, we also used lexicons introduced by Strapparava and Mihalcea [101], denoted as **EmoLex1**, and by Mohammad and Turney [80], denoted as **EmoLex2**. EmoLex1 contains six categories of words associated with one of Ekman's six basic emotions (i.e., happiness, sadness, fear, surprise, anger, and disgust), which is compiled from WordNet Affect [104]. Since WordNet Affect is built on top of the WordNet, it contains more general and formal emotional words. However, we need to use a lexicon to recognize emotional words that are usually used in social media and essentially do not appear in formal texts (e.g., bestfeeling). **EmoLex2** is the NRC Word-Emotion Association Lexicon that includes the collection of words that are captured from social media and categorized words based on Ekman's six basic emotions. EmoLex2 contains a large number of words in each category of emotions that are created by using a crowdsourcing approach in an iterative annotation fashion.

We use frequencies of subjective words, emotional words, and cancer-related keywords to construct lexicon-based feature vector in our computational models.

2.3.6. ConvLSTM

Prior studies [10, 113, 114] proposed computational models for analyzing emotional support for thousands of messages in OHCs by handcrafting a set of features extracted from patients' messages. Examples of such features include bag-of-words, linguistic features, lexicon-based features, and word patterns. Biyani et al. [2014] showed that the lexicons that contain a collection of subjective words, drug names, side effects, and therapeutic processes are among the most effective features in emotional vs. informational message classification.

[4] An example of online resource: https://www.cancer.net/navigating-cancer-care/side-effects

Name	Samples of words in lexicons
EmoLex1	eager, cheerful, sorrow, joyful, kindhearted, hilariously, gloom, heartbreak, regretful, panic, scare
EmoLex 2	superb, merrychristmas, bestfeeling, warms, healing, suffering, isolated, chronic, difficulties, injurious, crumbling
numWeak	conclusive, concrete, depression, dictator, enlighten, erase, fluent, floundering, imperfect, independence
numStrong	deploring, devoted, difficulty, enjoyable, equivocal, fondly, flourish, indefensible, mere, miraculously
numDrug	Cytosar-U, Cytoxan, Dabrafenib, Dacarbazine, Dacogen, Dactinomycin, Dasatinib, Daunorubicin Hydrochloride, Decitabine, Vectibix
numSide	osteoporosis, dysfunction, fibrosis, clot, pain, flash, stroke, fracture, palpitation, cognitive dysfunction, uterin cancer
numProc	radiology, chemo, chemotherapy, brachytherapy, mri, mammogram, mammosite, oncologist, oncology, aromata, inhibitor, lumpectomy

TABLE 2.5. Sample of emotional words compiled from Lexicons

However, despite the success of lexicon-based features, building comprehensive lexicons in different health domains (e.g., cancer and diabetes) requires medical experts on each domain to contribute intensively in a way that makes the process very challenging. For example, a group of experts should be employed to build several lexicons for breast cancer even though their productions have the least commonality with other diseases; moreover, all lexicons need to be frequently updated due to the fast advancement in the medical domain.

Besides, in OHCs, many messages provide emotional support without using explicit positive or negative affective words. For example, in this message from CSN: *"Keep in touch either on this site or e-mail direct anytime Just think of it as we are all strolling down the same path together and if one of us falls I know the rest will be right there to help out."*, the commentator presents an emotional support implicitly. Capturing these types of messages as emotional is chal-

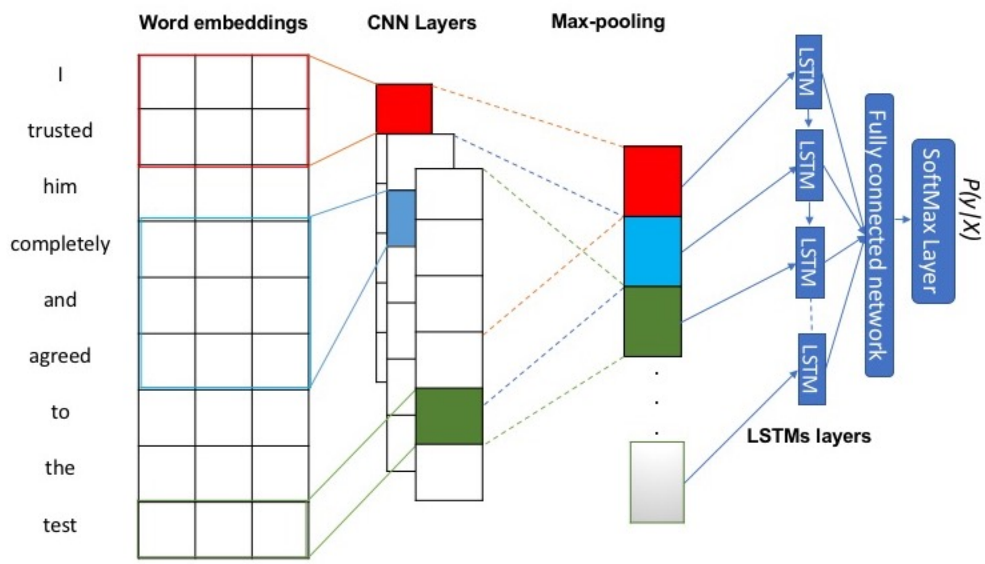

FIGURE 2.7. The structure of ConvLSTM.

lenging for traditional machine learning approaches which are mostly dependent upon lexical- or lexicon-based features.

In this section, we propose a computational model for analyzing *emotional messages* in OHCs that do not need to have handcrafted features or expensive lexicons, which require experts' knowledge, for generating high-quality results. Our proposed model called *deep convolutional LSTM*, or *ConvLSTM* for short (shown in Figure 2.7).

ConvLSTM is composed of convolutional neural networks (CNNs) and a one layer LSTM network. Extracting important features is one of the primary functions of CNNs [118]. CNNs generate semantic features from input messages by applying a convolutional function and using some kernels. The kernels are calibrated with the set of classes available for the task via a supervised training process that generates feature maps F whose units share the same weight and bias.

In ConvLSTM, the feature transfer task allows LSTM networks to represent and learn the function learned by a CNN network model. This model captures both high-level and sequential information without adding extra complexity like using the memory or attention mechanism [2, 106]. Regarding the architecture, our proposed classification model is close to the models stated in Kim et al. [2016] and Xiao and Cho [2016] where they applied a character-level CNN to create high-level features, whereas our model works at the word-level. We use the word-level input to use the benefit of applying embedding vectors which are trained on OHCs. We use the character-level model by Kim et al. [2016], denoted by C-ConvLSTM, as one strong baseline for comparison with our model.

As we mentioned, the input of an LSTM network at time t is the output of the CNN. Convolutional layers consist of two phases. In the first phase, k kernels for each region size (rs) are applied to the sequence of tokens in each message (e.g., 32 kernels for 2, 3, and 4 region sizes). Thus, the feature map $F \in \mathbb{R}^{k \times rs}$. The features f_t, with $t = 1, \cdots, T$, are defined as follows:

$$f_t = Relu(F[w_{t-(rs/2)+1}, \cdots, w_t, \cdots, w_{t+(rs/2)}]) \tag{10}$$

where *Relu* is the rectified linear unit activation function. This process is iteratively done for each time step (corresponding to each word) of the input message that ends up with $F = (f_1, f_2, \cdots, f_T)$ sequence. In the second phase, a max pooling function is applied to F that results in:

$$F' = (f_1{}', f'_2, ..., f'_{T/(poolingsize)})$$

F' is the output of the CNN that contains high-level, abstract features that are used as input to the subsequent LSTM network.

As Figure 2.7 depicts, the LSTM network generates a vector representation of each message, which is used by the Softmax function, $P(Y|W)$, for classification. Since both CNN network and LSTM network use a variety of hyper-parameters, we optimize parameters using a development set before applying them for identifying emotional messages and their emotion type.

As can be seen from the architecture of the ConvLSTM, this model does not include any contextual information that obviates the need for handcrafted features. However, we propose a variation of ConvLSTM which has a compositional structure and uses contextual information as the feature for the classification task.

2.3.7. ConvLexLSTM

Deep neural networks (DNNs), such as recurrent or convolutional neural network, can be fed with raw data (e.g., ConvLSTM). However, some studies showed that applying DNNs by using extracted features from raw data brings about better results [62, 86, 116].

Given a sentence of n words, we apply CNN to extract high-level (abstract) features that capture the semantic part of the text [62]. We combine high-level features with surface-level and lexicon-based features. Our proposed model, ConvLexLSTM, is shown graphically in Figure 4.1. As can be seen from the figure, we use a combination of CNN and LSTM models, where the final feature vectors from CNN augmented with lexicon-based features are fed as input to the LSTM network. The sequence of lexicon-based features in Figure 4.1 represents the best settings that achieved the highest result when applied to the development set. ConvLexLSTM captures all high-level, contextual and sequential information without leveraging extra parameters to the training process.

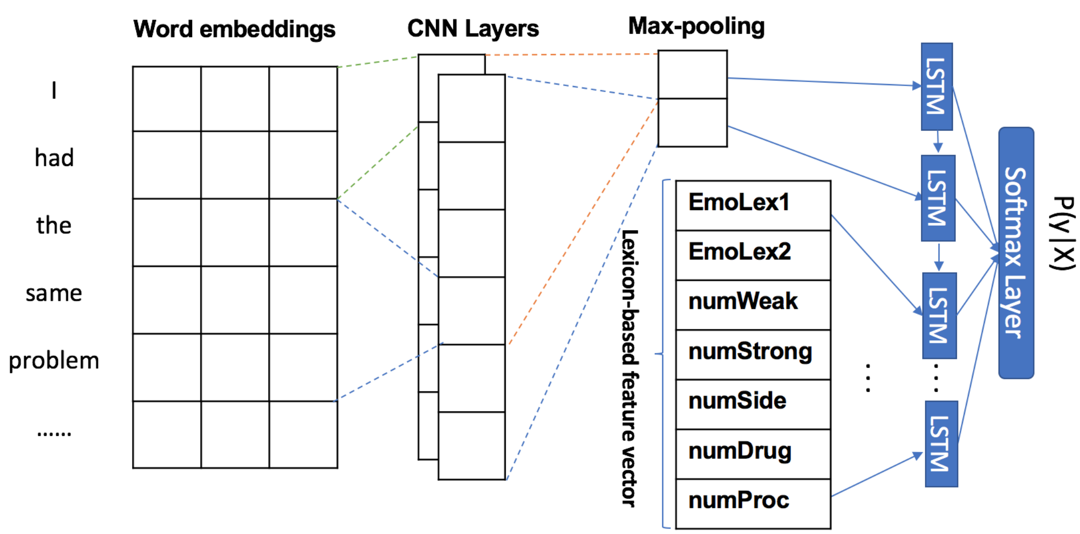

FIGURE 2.8. The structure of ConvLexLSTM.

2.4. Experimental Settings

We used our proposed models (i.e., ConvLSTM and ConvLexLSTM) for two classification tasks: emotional messages identification and emotion type detection in OHC messages. In this section, we also describe the evaluation of ConvLSTM and ConvLexLSTM. For the emotional messages identification task, we consider a binary classification which includes emotional/non-emotional classes. For evaluating these models for emotion type detection, we use the *joy* and *sadness* emotions, which have at least 5% coverage in our data (see Table 2). Also, since binary tasks are considered easier to learn than multi-class tasks [9], we trained our models in the two-class setting, joy/non-joy (and sad/non-sad), by binarizing the datasets.

In the experiments, for our word-level ConvLSTM, we used word embeddings as input to the neural networks. We used Wikipedia[5] and CSN (the dataset contains user's comments from June 2000 to June 2012) as our resources for generating word vectors by using the W2vector module in Gensim [92].

We estimated hyper-parameters for each model via a grid search over combinations of important parameters on a development set that consists of removing 20% of instances from the training set in each iteration of 10-fold cross-validation. We report precision, recall, and F1-score.

2.4.1. Hyper-parameter Settings

We set hyper-parameters for each DNN model (i.e., CNN, LSTM, C-ConvLSTM, ConvLSTM, and ConvLexLSTM) via grid search over combinations of important parameters except for word embedding vectors, which is explained in Section 2.3.4. We ran several experiments for optimizing each models' performance when applying them to the development set. We used the development set that is created for emotion type detection task and used the best settings of DNNs for both classification tasks (i.e., emotion identification and emotion type detection tasks). These parameters are learning rate (LR) [0.1,0.001], l_2 regularization (L2reg) $[0.0, 5E-5, 1E-5]$, decay rate $[0.0, 0.1, \cdots, 0.8]$, dropout $0.0, 0.1, \cdots, 0.6$, number of layers $[1, 2, \cdots, 10]$, pooling methods [1-Max, Mean, Last state] , order size in LSTM unigram, bigram, trigram, kernel region sizes (FRS) and number of feature maps (NF) in CNN (i.e., $[(2,3,4), (3,4,5), (4,5,6)]$, and $[32, 64, \cdots, 256]$). We set the hyper-parameters of Kim et al. [60] to the best setting provided by the authors.

As can be seen in Table 2.6, while LSTM network and CNN network achieved the best performance with the 3-layer setting, ConvLSTM and ConvLexLSTM obtained their best performance with a 1-layer setting; moreover, we can observe that the number of kernels (or filters) (NF) and figures in FRS in our proposed models and CNN are very different. Combining LSTM networks and CNN networks creates many learning parameters that need to be trained and learned during training phase. One observation is that by combining CNN networks and LSTM networks,

[5] https://dumps.wikimedia.org/enwiki/20160421

DNN models settings
–CNN: LR= 0.1, L2reg=1$E-5$, Decay rate=0.5, dropout=0.7, layer=3,
Max pooling, FRS=(2,3,4), NF=128
–LSTM: LR= 0.001, L2reg=1$E-5$, Decay rate=0.7, dropout=0.7, layer=3,
Max pooling
–ConvLSTM:LR= 0.001, L2reg=1$E-5$, Decay rate=0.7, dropout=0.6, layer=1,
Max pooling, FRS=(2,3), NF=32
–ConvLexLSTM:LR= 0.001, L2reg=1$E-5$, Decay rate=0.5, dropout=0.6, layer=2,
Max pooling, FRS=(2,3), NF=16

TABLE 2.6. Optimized hyperparameter settings for each model

we need more training data due to more parameters in the network need to be learned. However, by decreasing the number of FRS (from three regions in CNN network in two regions in ConvLSTM and ConvLexLSTM) and NF (from 128 in CNN network to 32 and 16 in ConvLSTM and ConvLexLSTM respectively) (see Table 2.6) the number of parameters in ConvLSTM and ConvLexLSTM plunges into almost the same number of parameters as individual CNN network and LSTM network. Besides, by decreasing the number of NF and FRS, we avoid possible overfitting caused by an overwhelming number of parameters and medium-sized datasets (B-DS and L-DS).

One of the most important hyper-parameters which has a significant impact on the performance of DNNs is the embedding vector, which has two parameters: source data used for training and the length of the vectors [126]. we created our word vectors in three different dimensions (i.e., 75, 150, 300) by using data compiled from Wikipedia and CSN as the training source. We create six different embedding vectors. To choose the best embedding vectors for our models, we trained ConvLSTM on the whole dataset (i.e., the remaining instances after separating the development set) and then computed the results when applying our model on the development set.

As we can see from Table 2.7, ConvLSTM with embedding vectors with the length of 150, which was created from CSN data, obtained the best results. The better performance of CSN-based embeddings in comparison with Wikipedia-based embeddings can be attributed to the higher

Type of resource & length	Precision(%)	Recall(%)	F1 (%)
csn-75	85.2	87.7	86.4
csn-150	**88.7**	**90.5**	**89.5**
csn-300	83.3	89.8	86.4
Wiki-75	80.9	87.6	84.1
Wiki-150	83.2	88.5	85.7
Wiki-300	82.0	88.2	84.9

TABLE 2.7. Results of the ConvLSTM on the development set using different embedding vectors

demands for random values initialization in a given message when Wikipedia resource embedding is applied. For example, considering the term *"bc"* which is an abbreviation term for *"breast cancer"*, *"bc"* has a pre-trained vector in the CSN-based embedding that represents a meaningful relation with any concepts related to the word *"cancer"*. However, the word *"bc"* does not exist in Wikipedia as a meaningful term, and ConvLSTM needs to initialize a random vector for the term *"bc"* when it appears in a message.

Among all hyperparameters, the length of the words' vector and type of resources had the highest impact on the F-1 measures. For example, changing vector size (which was trained on CSN) from 75 to 150 increased F1-score on the development sets as much as 3.1% (see Table 2.7). All models achieved their best results (i.e., emotional messages identification and emotion type detection) when assigned to the 150-dimension word embedding vectors trained on the CSN data.

We use our models with an optimized setting for identifying emotional messages and emotion type (i.e., joy, sadness) in OHCs' messages.

2.5. Performance of Our Proposed Models

We evaluate the performance of ConvLexLSTM and ConvLSTM on B-DS and L-DS for emotional messages identification and emotion type detection tasks.

Method	B-DS			L-DS		
	Precision (%)	**Recall (%)**	**F1 (%)**	**Precision (%)**	**Recall (%)**	**F1 (%)**
ConvLexLSTM	**91.8**	**96.6**	**94.1**	**84.3**	**87.0**	**85.6**
ConvLSTM	**88.7**	**94.3**	**91.4**	**82.7**	**84.5**	**83.6**
C-ConvLSTM	89.5	88.6	89.0	80.8	80.9	80.8
LSTM	88.4	91.5	89.9	82.3	79.8	81.0
CNN	84.6	92.8	88.5	80.1	81.3	80.7
EMO2014	85.1	91.1	88.0	74.0	83.2	78.3
BoW-POS	85.5	85.8	85.7	72.9	80.8	76.6
Lexicon-based model	66.7	92.0	77.4	64.3	97.3	77.4

TABLE 2.8. Emotional messages classification results using 10-fold cross validation.

2.5.1. Results for Emotional Messages Identification

From Table 2.8, we observe that all deep neural networks achieve a better performance than the state-of-the-art model, EMO2014 by Biyani et al. [2014] and the BoW baseline and the lexicon-based model. ConvLexLSTM achieves the best performance with an F1-score of 94.1% and 85.6% on B-DS and L-DS, respectively. ConvLexLSTM outperforms EMO2014 [10] by 6.1% and 7.3% on the B-DS and L-DS, respectively. Also, it can be seen that the word-level ConvLexLSTM and ConvLSTM outperform the character-level model of Kim et al. [2016] (C-ConvLSTM), which supports our hypothesis that using word embeddings trained on OHC health data yields better model performance than character-level. As we can see from Table 2.8, combining lexicons with the extracted feature from CNN improved the results (comparing the results of ConvLSTM and ConvLexLSTM). We also can observe that ConvLSTM which does not use any contextual information and can be applied on any type of texts, achieved better F1-scores than the state-of-the-art model (i.e., EMO2014) by 2.4% on B-DS and 3.6% on L-DS dataset. High-quality results obtained by ConvLSTM shows that it can be applied in any OHC forums that handcrafting contextual feature is expensive or challenging; In cases where contextual features are provided, we show that ConLexLSTM improves the quality of the classification model.

The results presented in Table 2.8 show that all classifiers achieve substantially better Precision, Recall and F1-score on B-DS compared with L-DS. The F1-score decline in L-DS can be explained by the fact that in B-DS, commentators express their emotions using more explicit emotional words compared to L-DS; for example, consider the messages: *"I haven't gotten depressed from it"* and *"I have no idea what the future holds."* In the former message from B-DS, the commentator used the word *depressed* reflecting a rather similar situation (this judgment is based on reading the whole related threads) as in L-DS, but the commentator in L-DS did not use any explicit emotional words. This is an interesting observation since, intuitively, women typically use more emotional words as compared to men. A manual investigation of our B-DS and L-DS datasets (based solely on the text of the messages) revealed that 79% of messages were written by women in B-DS, whereas only 55% of the messages were written by women in L-DS. We found that even men were more emotional when writing in the B-DS in comparison with L-DS, such that 83.8% of messages written by men in the B-DS were emotional whereas only 66.3% in the L-DS.[6]

2.5.2. Results for Emotion Type Detection

Due to the complexity of the classification task in emotion type detection such that the classifier should identify emotional messages and then detect the type of emotion in messages, we decide to perform ablation experiment starting with ConvLexLSTM. First, we evaluate ConvLexLSTM performance in an ablation experiment to determine the role played by each component for emotion detection. Specifically, we compare ConvLexLSTM with ConvLSTM (a model that has the same architecture as ConvLexLSTM but does not use any external lexicon), CNN, LSTM, and support vector machine (SVM) with the (concatenated) features from the seven lexicons (described above).

Table 2.9 shows the results of this comparison. As can be seen from the table, ConvLexLSTM achieves the best results consistently throughout all experiments in terms of all compared measures. This ablation experiment confirms our intuition that all components are contributing to the final emotion detection, even if some of them to a smaller extent. For example, removing the seven lexicon features from ConvLexLSTM, which yields ConvLSTM, results in a drop in F1-score by 5.8% and 4.8% on *joy* in B-DS and L-DS, respectively, and 3.9% and 5.4% on *sadness* in

[6]Gender distributions in B-DS were 84 (F) and 28 (M)

B-DS and L-DS, respectively.

Not surprisingly, SVM with the seven-lexicon based features (denoted as Seven-Lexicon) performs the worst among the compared models, suggesting that capturing the semantic information from text improves emotion detection. Still, ConvLSTM is the second best performing model in terms of F1-score. These results show that our model can be successfully applied in a health domain even in the absence of health lexicons, which are often expensive to obtain. Not surprisingly, the SVM with the seven-lexicon based features (denoted as Seven-Lexicon) performs the worst among the compared models, suggesting that capturing the semantic information from text improves emotion detection.

Second, we compare ConvLexLSTM with three baselines:

(1) C-ConvLSTM which is a character-level CNN-LSTM) proposed by Kim et al. [2016]

(2) SWAT [54], which is an emotion detection model from SemEval-2007. We re-implement SWAT with a minor change: using WordNet instead of Roger's Thesaurus for finding unigrams' synonyms.

(3) EmoSVM, which is an SVM with a set of handcrafted features. We developed this model to represent traditional style of classification. We elaborate on the applied in the following.

In developing EmoSVM, we used the following features as input to the machine learning algorithms:

- **unigrams and bigrams:** n-grams have been shown to improve classification results in social media analysis tasks [117]. We treat each message drawn from our dataset as a document and then calculate normalized *tf-idf* for each unigram and bigram in the vocabulary. We performed stemming and eliminated punctuation from each message.
- **Part-of-speech (POS):** POS tags obtain primary properties of each word in OHC messages. Using POS tags as a feature for the classification task in sentiment and subjectivity task has been a common practice [10, 120]. We encode the frequency of POS tags in a message as a feature in developing EmoSVM.
- **Emotional words:** Intuitively, people use emotional words when they want to express

	Method	B-DS			L-DS		
		Precision (%)	**Recall (%)**	**F1 (%)**	**Precision (%)**	**Recall (%)**	**F1 (%)**
Joy	ConvLexLSTM	**92.3**	**94.3**	**93.2**	**90.4**	**89.3**	**89.8**
	ConvLSTM	86.6	88.4	87.4	87.0	83.0	85.0
	CNN	85.0	84.0	84.5	82.2	82.8	82.5
	LSTM	86.0	86.6	86.3	85.0	83.0	84.0
	Seven-Lexicon	63.4	87.3	73.45	60.0	85.1	70.37
	C-ConvLSTM	86.2	87.0	86.6	85.0	82.0	83.47
	SWAT	66.0	68.0	67.0	65.5	66.7	66.0
	EmoSVM	81.0	82.0	81.5	82.0	80.0	81.0
Sad	ConvLexLSTM	**93.7**	**91.1**	**92.3**	**88.0**	**90.9**	**89.4**
	ConvLSTM	89.0	87.8	88.4	81.0	87.5	84.0
	CNN	83.2	83.6	83.4	81.7	80.5	81.0
	LSTM	87.4	85.8	86.6	83.2	83.6	83.4
	Seven-Lexicon	61.0	84.9	70.99	61.0	83.3	70.42
	C-ConvLSTM	85.0	83.6	84.3	83.7	82.1	82.9
	SWAT	65.0	66.0	65.5	64.0	65.0	64.5
	EmoSVM	80.5	81.7	81.0	79.0	78.0	78.5

TABLE 2.9. Emotion detection results using 10-fold cross validation. The numbers are percentages.

their feelings. Building and using lexicons for identifying emotional sentences has been used extensively in emotion detection tasks [79, 80, 103, 104]. We used the word-emotions association lexicon by Mohammad [79] and the WordNet-Affect lexicon by Strapparava et al. [104] to build two binary features that show the existence of any type of emotional word in the sentence. For example, if the word *excited* exist in a sentence for the joy vs. non-joy classification task, then we encode this feature as *"1"* in our classification model.

- **Sentiments of sentence:** Intuitively, it is more likely that a person who has a positive opinion towards something, he/she will express a positive emotion like *joy* rather than *sad, disgust, fear* emotions and vice versa. We applied the output of the Stanford sentiment tool developed by Socher et al. [100] as a feature to show the positivity or negativity of a sentence. We encode this feature as *1 , 0, and -1* for the positive, neutral, and negative outputs of the tool, respectively.

We used the development set and performed an ablation study to confirm that all of the above features have a positive impact on the results. We kept the above features in EmoSVM model and discard the ones which had a negative impact on the results. Table 2.9 shows the results of comparing the performance of our baselines with our proposed models. As seen from the table, ConvLexLSTM outperforms all three baselines, and more importantly, the character-level CNN-LSTM by Kim et al. [60] (i.e., the C-ConvLSTM model). We use the word-level input to use the benefit of applying embedding vectors which are trained on OHCs and using lexicon-based features. We use the word-level input in our models to gain the benefit of applying embedding vectors which are trained on CSN. This result confirms our belief that applying embedding vectors, which are trained directly on data from OHCs, yields improvement in performance over character-level models.

It is worth mentioning that all deep neural networks, ConvLexLSTM, ConvLSTM, CNN, LSTM, and C-ConvLSTM, that capture high-level semantic features perform better than the traditional models of emotion detection. The lexicon-based features act as a complement (for the high-level semantic features) by looking into exact words in the text to generate appropriate features in ConvLexLSTM for emotion detection. With a paired T-test, the improvements of ConvLexLSTM over the compared models for F1-score are statistically significant for p-values < 0.05.

2.5.3. Impact of Holidays on Emotional States

We further analyzed the impact of several US holidays (shown in Figure 2.9) on CSN users' emotional states. For this experiment, we extracted messages from two days before and two days after each holiday. We also created a dataset by combining messages written on five random days, i.e., not close to any event, to be used as a baseline for comparing the emotional state of participants

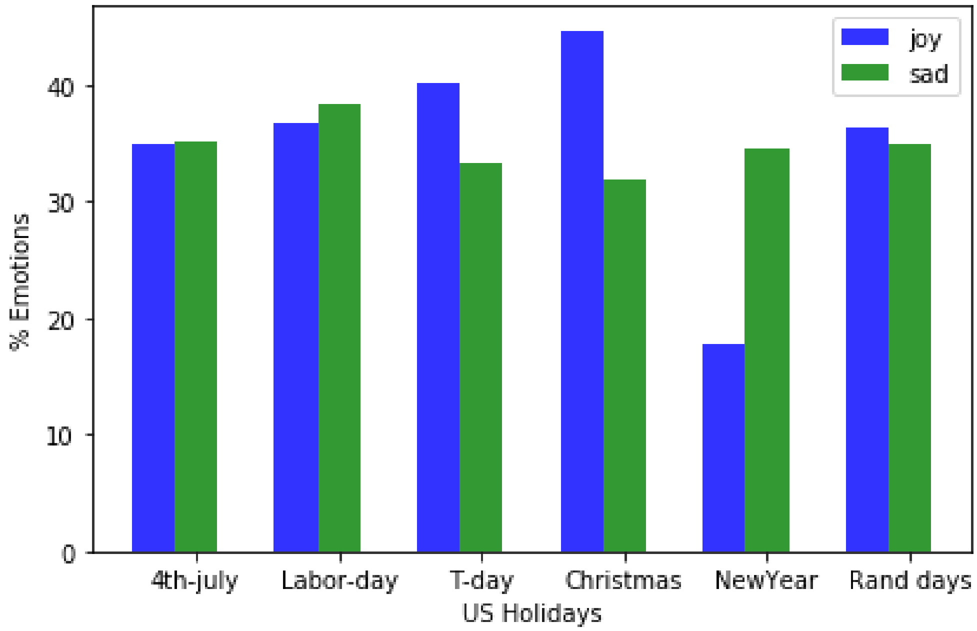

FIGURE 2.9. Joy and sadness throughout events.

in different holidays. We used ConvLexLSTM to classify joy and sadness on these data. As can be seen from Figure 2.9, on a typical day, the percentage of joy and sad emotions are similar, whereas Christmas and Thanksgiving show more joyful spirits, possibly due to family gatherings and other social events around these holidays, in which people feel supported and hence feel better. This is an interesting direction of investigation for future work.

2.6. Emotion Type Distribution in Influential Users Posts

OHCs have a group of participants, like other social networks (e.g., Facebook, Twitter, etc.), who have high influence on their peers in the network; these participants are called influential users, who are implicitly the leader of the network. In OHCs, the importance of influential users is to the extent that some studies tailor the effectiveness of OHCs to the influential users' activities and the population of them in the network [75]. Furthermore, influential users have the ability to impact the the whole OHC network implicitly and indirectly [81]. Prior studies postulated that health-related behaviors and the emotional state of patients are highly influenced by the quality of social contact [16, 18, 36]; this shows the potential of influential users in changing patients' behavior, given the high connections that they make in OHCs [10].

Given the importance of the roles that influential users play in OHCs, identifying and tracking their behaviors is one of the highest priorities of OHC moderators [16]. For example, moderators can ask influential users to help patients who express signs of depression in the network (e.g., patients who post many messages with sad content); or moderators can use the communicative power of influential users to encourage patients in using newly established therapeutic procedures, correcting wrong beliefs about a kind of therapy and side-effects, or encouraging patients to have a healthy lifestyle [14, 108].

In OHCs, moderators need to detect influential users who spread positive emotions in the network and connect them to depressed patients for the sake of encouraging them emotionally. We used the list of user-ids that is provided by CSN moderators [10], which include 62 influential users in the breast cancer forum. We used the CSN data that was collected from June 2000 until June 2012 and extracted all messages that were posted by influential users, which are 286,487 sentences in total. We eliminated all sentences that we used in building our dataset (i.e., B-DS)

We also randomly chose a list of 6 influential users and created an individual dataset for each influential user from their messages that were posted from June 2000 until June 2011. We call data for each influential user $inf-1, inf-2, \cdots, inf-6$.

First, We applied our computational model for emotional messages identification on each of six influential users data and the whole data of all influential users (see Figure 2.10).

As we can see from Figure 2.10, while emotional messages in individual's posts are always more that non-emotional posts, the number of emotional and non-emotional posts are almost the same in the whole influential users' posts. From Figure 2.10, we can see that in four of the six influential users, they are providing almost the same amount of emotional and non-emotional messages; however, in two of them (i.e., inf-3 and inf-4), they tend to support patients emotionally considerably more than other influential users.

We then apply our emotion type detection model on the emotional message that was extracted from the previous experiment. Figure 2.11 illustrates the difference between messages with joy content and messages with sadness content. As can be seen from Figure 2.11, the percentage of messages with joy content is considerably more than the messages with sad content. These re-

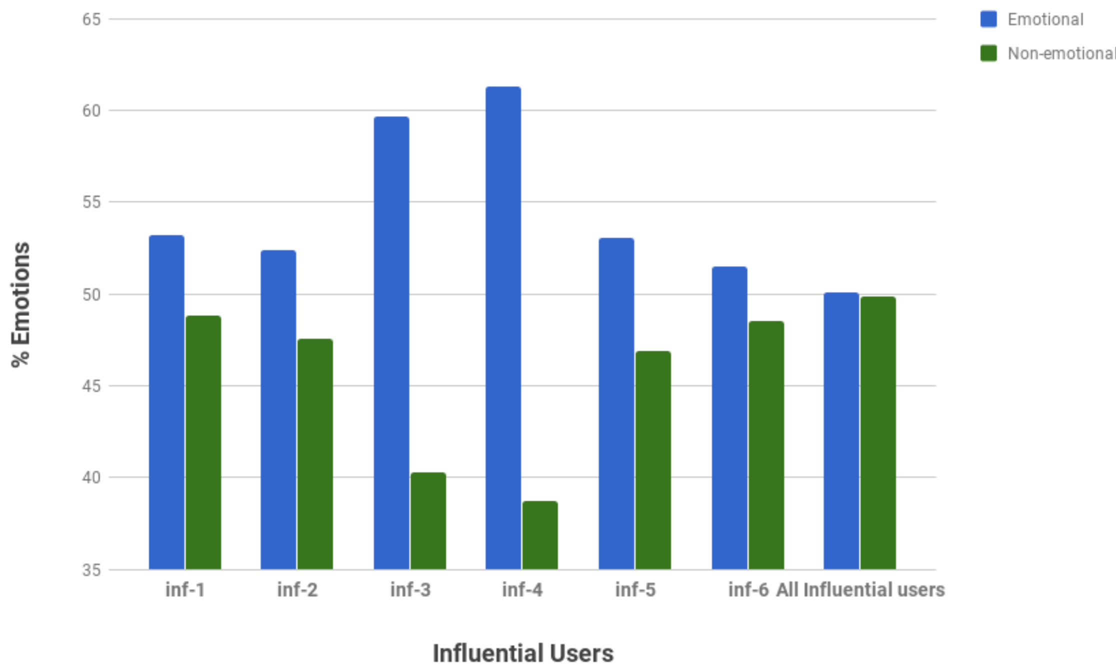

FIGURE 2.10. Comparing emotional and non-emotional sentences in influential users' posts

sults show that influential users tend to encourage patient with emotionally joy content rather than sad content. We explored through messages with sad content and found that many of them are a sympathetic message that an influential user intended to show his/her feelings of sorrow for other patient.

The experiments in this section show that our proposed models can be used for tracking emotional content in influential users' messages and help moderators in finding them and connecting them to patients who are suffering from depression or deep sadness.

2.7. Related Work

Emotion detection has been studied in computational linguistics for a long time. Identifying emotions latent in OHCs' messages is one of the primary steps in analyzing social support. The most popular model for emotion classification is based on the Ekman's basic emotion set [104]. This set includes anger, disgust, fear, happiness, sadness, and surprise, and is arguably the best well-known emotion categorization [32].

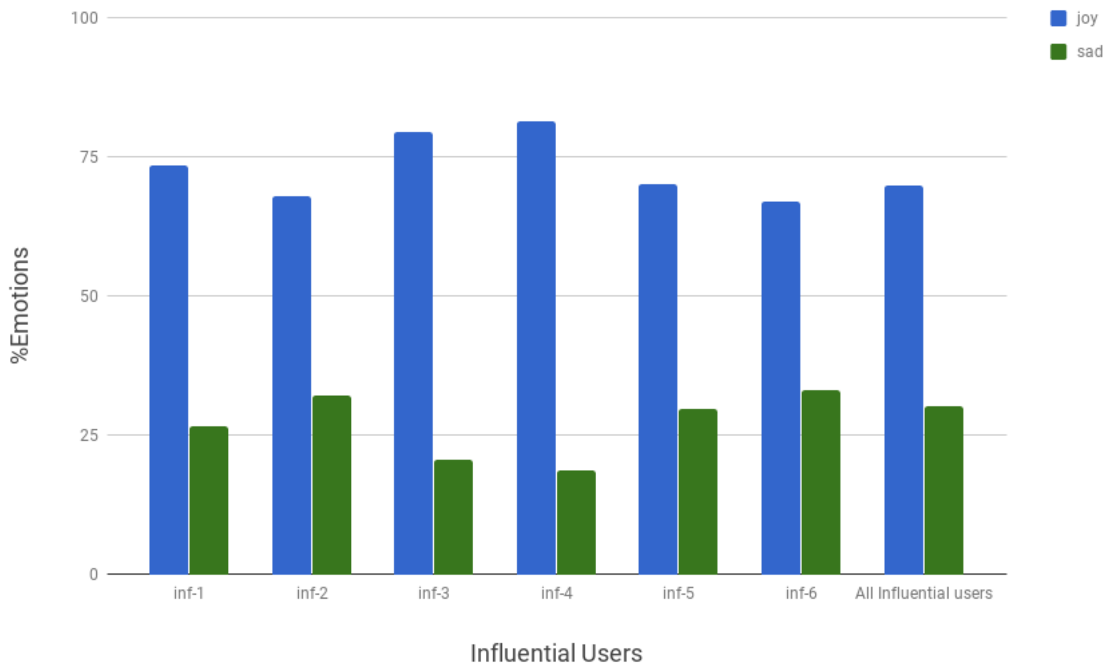

FIGURE 2.11. Comparing sentences with joy and sadness emotions in influential users' posts

Strapparava and Mihalcea [102] proposed knowledge-based and corpus-based methods for classifying emotions based on Ekman's basic emotions. Co-occurrence distribution of the general words in the text with emotional words have been used by Katz et al. [54] for identifying emotion type latent in news headlines. Keyword-based approaches, which are based on finding emotional words in the text, suffer from word meaning ambiguity that causes the inability to classify texts that lacks specific keywords. Therefore, Bao et al. [4] considered topical relationships across words and emotion type and Agrawal and An [1] proposed semantic relations between words and the emotion type for classifying emotions.

Emotion detection approaches in social media have different specialties due to their informal context in which people do not follow grammatical rules and use many characters that do not occur in formal texts (e.g. #, :)). Some works created emotion lexicons for social media and showed that they improve the accuracy of their models [79, 105]. For example, Mohammad [79] proposed a method for creating a lexicon based on *emotion word hashtags* and showed that using this lexicon improves the results.

Although emotion detection from social media has been studied for years, there are only a few studies that applied computational models on OHCs for detecting emotions [10, 112, 113, 114]. Wang et al. [114] used a linear regression model to identify emotional supports in messages from a cancer forum. For a given message, the trained model predicts to what extent each sentence contains either emotional or non-emotional supports on a scale of 1 to 7. Three types of features for classification were created: psycholinguistics features (e.g., LIWC), linguistic features (e.g., POS, messages type, message length, subjectivity intensity, etc.), and LDA topical features (e.g., family and friends, diet, etc.). The authors showed that linguistic and topical features were the most effective ones.

Since messages posted in OHCs may contain a combination of emotional and non-emotional support, Biyani et al. [10] performed emotional message classification at the sentence level. The authors annotated $1,000+$ sentences, which were extracted from random posts in the breast cancer discussion board of the cancer survivor network (CSN[7]). For identifying emotional messages, the authors used unigrams, POS tags, structural patterns, and five different lexicons that contain strong and weak subjective words, cancer drugs, side-effects, and cancer procedures. Biyani et al. [10] showed that features drawn from lexicons had a high impact on the results. Wang et al. [113] classified OHCs' messages based on the intention of the participant when writing messages (i.e., companionship, seeking information, seeking emotion, providing information, and providing emotion). The authors applied four types of features, three of which were used by Wang et al. [114] coupled with lexicon-based features used in Biyani et al. [10].

The above studies focused on classifying emotional messages or commentators' intentions, using hand-crafted features. In contrast, we improve these works by using state-of-the-art methods (deep neural networks) for emotional message classification and provide a more insightful understanding of emotional messages by identifying their emotion types. Despite the large use of deep learning approaches in the biomedical domain, including electronic health records and imaging [78], deep learning was not extensively used for analyzing short messages posted in online health communities.

[7]https://csn.cancer.org

2.8. Chapter Summary and Future Directions

In this chapter, we addressed the problem of emotional messages identification and emotion type detection from messages that are posted in online health communities. To this end, we proposed a computational model, ConvLexLSTM, that combines the strengths of CNN networks, LSTM networks and lexicon-based approaches to capture the hidden semantics in OHCs messages. We show that our proposed model, with or without lexicon-based features, which are often expensive to obtain or maintain in a health domain, provides a better emotion type detection compared with strong baselines and prior works.

We showed two benefits of using our models as a feature in OHC applications. The results of our models can be used by OHC moderators for manging purposes and researchers in the medical and psychological domain. First we applied our model for emotion type detection around important holidays in USA and secondly, we used our model to analyze influential users' behavior when interacting with patients.

In the future, it would be interesting to extend this work to other types of emotions, e.g., anger or fear, as well as to other OHCs. Since emotion type detection has not been used in identifying influential users, it will be interesting to use emotions in influential users' posts as a feature in combination with other features, which have been studied before, and measure the impact of emotion-based features in identifying influential users. Besides, investigating the relation between posting messages with sad content and users' depression is worth to be studied.

CHAPTER 3

ANALYZING EMPATHETIC MESSAGES IN ONLINE HEALTH COMMUNITIES

[1] Empathy captures one's ability to correlate with and understand others' emotional states and experiences. Messages with empathetic content are considered as one of the main advantages for joining online health communities due to their potential to improve people's moods. Unfortunately, to this date, no computational studies exist that automatically identify empathetic messages in online health communities and analyze its impact on people's emotional states. In this chapter, we propose a combination of Convolutional Neural Networks (CNNs) and Long Short Term Memory (LSTMs) networks, and show that the proposed model outperforms each individual model (CNN and LSTM) as well as several baselines.

3.1. Introduction

Empathy captures the ability of an individual to correlate with and gain an accurate understanding of other individuals' emotional states by putting oneself in their situations with appropriate reactions [6, 64]. Empathy is shown to have a fundamental role in connecting people in a community together [27]. Recently, many studies in social and psychological sciences have investigated the correlation between the empathetic capability of users in a social network and their characteristics and behavioral patterns. For example, Kardos et al. [2017] analyzed social networks and found that higher empathetic abilities in social network users bring about a bigger size of close friends' list and vice versa. Medeiros and Bosse [2016] and Coursaris and Liu [2009] also expressed that empathetic abilities account for social supports in social media, and Mayshak et al. [2017] showed that the level of user engagement with social networking website has a direct correlation with empathetic abilities. Finally, Del Rey et al. [2016] suggested that empathy negatively predicts traditional bullying and cyber-bullying perpetration.

Recent studies show that in the health domain, empathy is one of the main advantages of

[1]Appreciation is extended to association for computational linguistics (ACL). This entire chapter is reproduced from Khanpour, H., Caragea, C. and Biyani, P. [2017], Identifying empathetic messages in online health communities, in Proceedings of the Eighth International Joint Conference on Natural Language Processing, Vol. 2, pp. 246251, with permission from ACL

–***Patient:*** *Hi all sense being on chemo (5 down 1 to tch) with the last two really I have had a problem with my BP being high. I am having a problem with my heart racing. At rest it may get down to 86. When my oncologist did the muga scan it went from 68 to 63. I have never had a problem with my heart at all. I'm Very nervous.*
–***Commentator***: I had much the same problem while doing chemo, the last 2 or 3 rounds were the worst. Try not to worry to much! By the way I am the proud owner of 3 chihuahuas. Blessings to you...Alison
–***Patient:*** *Thanks so much I feel allot better now. I did talk to my Dr and he is giving me meds to lower the rate. I feel like I spend my time fighting side affects LOL. Thanks sisters. Take care all*

TABLE 3.1. A sample of an empathetic message and its impact on patient's emotion

using online health communities (OHCs) [70, 74, 82], which potentially foster the healing process by decreasing distress and increasing optimism [42, 85]. Table 3.1 shows an anecdotal example, extracted from a cancer-related community, illustrating the function of empathetic messages.

Despite its importance, to our knowledge, there has not been any computational approaches proposed for identifying and analyzing empathetic messages in OHCs. The works above, in social science and psychology, are based upon using questionnaire, direct interviews, or at most hundreds of samples from manually collected data. These studies suffer from scalability, biased data usage [89], and high dependency to the human's memory that might not recall details accurately [66, 91].

In this chapter, we propose a computational approach, which is able to analyze large numbers of messages in social networks and automatically identify messages that contain empathy to make an appropriate foundation for further, deeper, and scalable studies and developing applications. Automatic empathetic message identifier can be used by social networks moderators for monitoring communities mental health, cyber-bullying and cyber-stalking detection, measuring the level of users engagement in communities [73], predicting users' position in online communities [53], as well as the loneliness of users [87]. Furthermore, such an application can be employed in measuring nursing skills [121], measuring the quality of online counseling sessions, and assessing the quality of human-robot interactions [40, 65].

The scope of this study is on analyzing patients' comments in OHCs to identify empathetic messages (i.e., sentences in comments). While providing patients' with empathetic messages on OHCs is a critical outcome of online communities, computational studies addressed only two general functions of OHCs: informational seeking forum and social support forum [23, 34]. Recently, some computational models for analyzing social supports [10, 89, 113] and identifying influential users [127] have been proposed, which are the closest studies in the literature to our work. However, in our study, we are addressing a different problem, that of identifying empathetic messages.

Our contributions in this work are as follows:

(1) We create a dataset annotated for empathetic messages identification in OHCs. This is the first study which provides a dataset for identifying empathetic messages in social media.
(2) We propose an automatic model for identifying empathetic messages in OHCs. To the best of our knowledge, this is the first work that proposes a computational model for detecting empathy in the health-related social networks.
(3) We experimentally validate our empathy identification model on a manually annotated dataset specifically created for this task from a Cancer Survivors' Network.
(4) We also show that generally empathetic messages in comments of OHCs transform participants feelings from negative to positive.

3.2. Data Collection and Annotation

We randomly selected 225 comments from 21 discussion threads in a lung cancer discussion board in a cancer survivor network (CSN)[2]. Following Biyani et al. [2014], we selected messages (i.e., sentences in comments) with length greater than four words. We ended up with 1041 messages in total. We integrated our collected data with 1066 messages extracted from the breast cancer discussion board in CSN that was provided by Biyani et al. [2014].

The purpose of the annotation is to identify empathetic messages by which a commentator intended to show his/her empathy with other people. Two annotators (graduate students) contributed to the annotation task. We asked them to update their knowledge of empathy by reading two studies (i.e., Collins [2014] and Decety and Jackson [2004]) during a week. After a group meeting between annotators and researchers to share and discuss their understanding of empathetic

[2]https://csn.cancer.org

messages with the presence of two psychologists, the annotation task began in an iterative fashion similar to prior studies and guidelines [31, 35, 97] . In each round, 200 messages were assigned and annotators were asked to meet with researchers to discuss disagreements. 100% inter-annotator agreement (IAA) were achieved after each round of discussions. After three initial rounds of annotations where IAA exceeded 80%, the remaining of the data were assigned (1417 messages) to the annotators where they achieved 87% IAA. The last round of assigned data was adjudicated by one of the main researchers. Table 3.2 represents the distribution of empathetic messages in each datasets (i.e., breast cancer (B-dataset) and lung cancer (L-dataset)). As can be seen, in B-dataset the percentage of the empathetic messages is more than 1.5 times of the L-dataset.

Dataset	**Empathetic msgs.**	**Percentage(%)**
B-dataset	494 out of 1066	46.3
L-dataset	295 out of 1041	28.3

TABLE 3.2. Statistics from the data collections.

3.3. Model

In this section, we describe our proposed model for empathetic messages identification in OHCs.

Given a message in a social network, $S = \{W_1, W_2, \cdots, W_i, \cdots, W_n\}$ with n words, empathy classifier aims at determining if the message contains empathetic content.

Hyper-parameter Settings
–LSTM: W2vec-S=150, LR= 0.001, L2reg=$1E-5$, Decay rate=0.7, dropout=0.5, layer=5, Max pooling
–ConvLST:W2vec-S=150, LR= 0.01, L2reg=$1E-5$, Decay rate=0.7, dropout=0.5, layer=2, Max pooling
–CNN: W2vec-S=150, LR= 0.1, L2reg=$1E-5$, Decay rate=0.5, dropout=0.5, layer=2, Max pooling, FRS=(1,2,3), NF=64

TABLE 3.3. Hyperparameter settings for each model

3.3.1. Word Representations

We use word embedding with an embedding matrix $E_w \in R^{d_w \times V_w}$ where d_w is the embedding dimension and V_m shows word vocabulary size. We generate six word embedding matrices by using two resources: Wikipedia [3] and CSN (the whole CSN data, 752.4 MB) in three different dimensions (i.e., 75, 150, and 300). We used W2vector module in Gensim [92].

3.3.2. Model Description

The proposed model for classifying empathetic messages is a combination of a convolutional and an LSTM network (ConvLST). This network takes word embeddings as input and creates a sequence of dense, real-valued vectors: $E = (e_1, e_2, \cdots, e_T)$. By applying multiple convolutional layers to E and using pooling, we obtain a dynamic sequence of feature vectors: $F = (f_1, f_2, \cdots, f_n)$ which is fed into a variation of recurrent neural network (LSTM). The output of the LSTM network is given to a softmax function to compute the predictive probabilities, $p(y = k|S)$, of each binary classes given message S (see Figure 3.1).

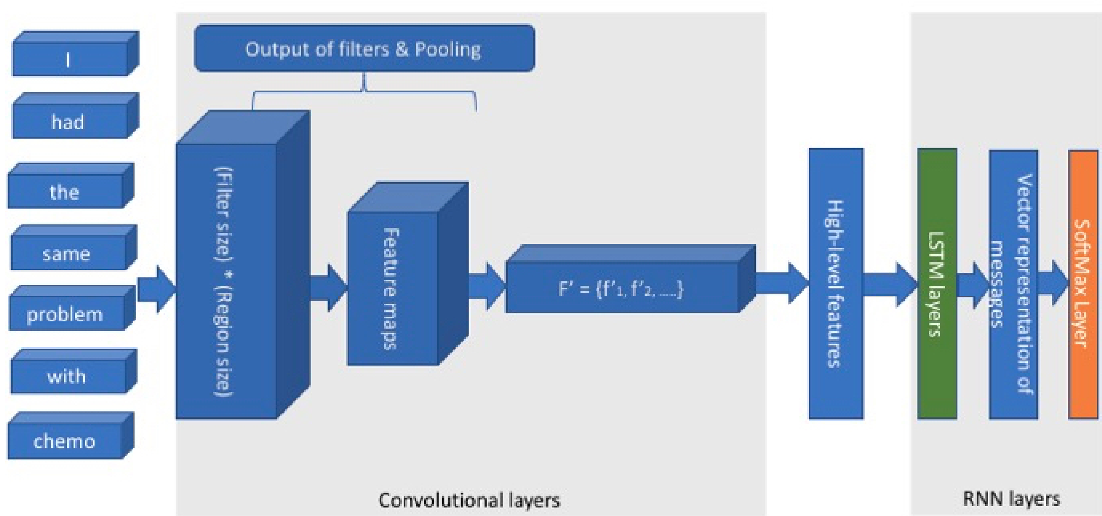

FIGURE 3.1. ConvLST structure for empathetic messages identification.

[3] https://dumps.wikimedia.org/enwiki/20160421

3.4. Experiments

In this sections, we present our optimization process and the results of our model used to identify empathetic messages.

3.4.1. Hyper-parameters settings

We separated 15% of the dataset as the developing set. We optimized hyperparameters of ConvLST and each of embedded models (i.e., CNN and LSTM) to compare their performances with ConvLST. We obtain the best optimization settings for hyperparameters via grid search over combinations of important parameters. These parameters are: word embedding vector size (W2vec-S) $[75, 150, 300]$,initial learning rate (LR) [0.1,0.001], l_2 regularization (L2reg) $[0.0, 5E-5, 1E-5]$, decay rate $[0.0, 0.1, \cdots, 0.8]$, dropout $0.0, 0.1, \cdots, 0.6$, number of layers $[1, 2, \cdots, 10]$, pooling methods [1-Max, Mean, Last state], order size in LSTM unigram, bigram, trigram, filter region sizes (FRS) and number of feature maps (NF) in CNN, $[(1,2,3),(2,3,4),(3,4,5),(4,5,6)]$ and $[32, 64, \cdots, 256]$, respectively.

Baselines: We compare our models with the following baselines:

(1) **Bag-of-words and POS tags:** Word frequencies and their part-of-speech tags show the primary property of the text and has been used in studies on OHCs' message processing [10]. We used both words and their POS tags' frequencies as features.

(2) **Lexicon-based model:** Lexicon-based approaches has been used in many studies related to emotion detection [101, 104] and sentiment analysis tasks [67, 79]. Following Biyani et al. [2014] who used lexical features, we used the same lexica, which are provided by Biyani et al. [2014]. These lexica include: weak subjective words, strong subjective words, cancer drugs, side-effects, and therapeutic procedures, for building our baseline's feature set.

3.4.2. Empathetic Message Identification

Table 3.4 compares the performance of our proposed model (ConvLST) with other DNNs (CNN and LSTM), and the baselines. As can be seen, our model achieves the best F1-score, which is 8.46% higher than the best baseline's F1-score (i.e., 69.90%). These results show that the

Method	**P(%)**	**R(%)**	**F-1 (%)**
ConvLST	78.61	**78.12**	**78.36**
LSTM	**79.47**	75.00	77.17
CNN	76.20	77.00	76.60
BoW+POS	71.8	68.2	69.90
Lexical-based	54.5	46.9	50.4

TABLE 3.4. Empathetic message identification.

combination of CNN and LSTM (ConvLST), that employs the sequences of important features extracted by CNNs achieves a better performance compared with each of CNNs and LSTMs models individually.

While LSTM network achieved the best precision score across all models, ConvLST gained the highest F1-score and recall, among all models. Table 3.4 also shows that Lexical-based baseline resulted in the least F1-score. Lexicon-based baseline uses two categories of features: subjectivity-related and informational-related features. By removing subjectivity-related features, F1-score drops to 15.7% and by eliminating informational-related features, F1-score drops to 47.3%. These results suggest that the subjectivity features are more effective than the informational-related ones.

3.4.3. Sentiment Dynamics with Empathetic Messages

In this section, we conduct an experiment to investigate the potential of empathetic messages for changing a thread originator's feelings. We used the data extracted from CSN, which include users' comments from June 2000 to June 2012. We extracted all threads where the originator of a thread replied once after at least one empathetic comment were posted from other users (responders). We followed the same experimental setting presented in Qiu et al. [2011]. In total, 12915 threads were extracted for analysis, which include 250,868 comments posted by 5516 users.

We ran our classifier for empathetic message identification on all responders' messages which were posted between two posts of the originator (e.g., ***Commentator***'s post in Table 3.1). We also discarded messages in which an initiator simply thanks a fellow member and used a threshold of four on the number of words [10]. We ran Stanford sentiment toolkit [72] on the originator'

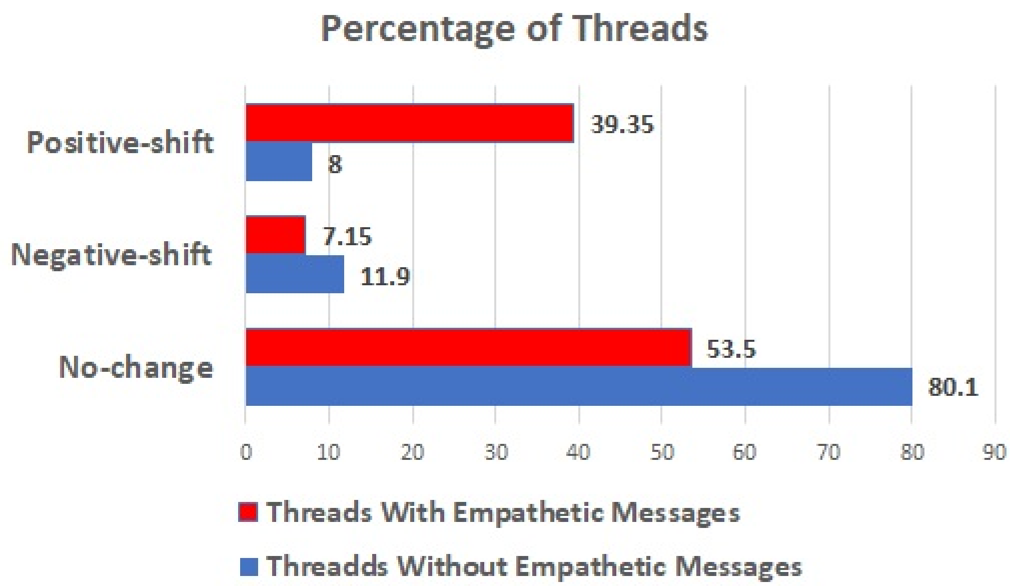

FIGURE 3.2. Thread-initiator's feeling transformation as a result of empathetic messages in a thread.

posts (e.g., ***Patient***'s posts in Table 3.1). In this way, it is possible to determine whether the empathetic messages provided by responders who replied to the thread, are able to change the sentiment of the thread originator. To better understand any changes in feelings, we categorized changes in three groups, i.e., Positive-shift, Negative-shift, and No-change. Positive-shift represents any transformation from worse to better feelings such as negative-to-positive, neutral-to-positive, negative-to-neutral. Negative-shift has a converse settings compared with the Positive-shift and No-change represents a state that originator's second post reflects the same sentiment as the initial one.

These results are shown in Figure 3.2 (the red bars). As can be seen from the figure, in 39.35% of the threads, empathetic messages bring a positive-shift in originators' feelings as opposed to only 7.15% negative-shift. We can also observe that in 53.5% of the threads, the originators' feelings do not change. Thus, we can conclude that empathetic messages play a major role in improving participants' feelings in OHCs.

We also contrasted the positive-shift, negative-shift, and no-change in the threads with empathetic messages (the red bars in Figure 3.2) with those in the threads without empathetic messages (the blue bars in Figure 3.2) to better understand the impact of empathy on people's moods.

More precisely, we ran the sentiment tool over the threads with no empathetic messages and found that only 8% positive shift, 11.9% negative shift and 80.1% no-change occurred. These results suggest that positive sentiment changes occur more prominently in threads containing empathetic messages compared to those with no empathetic messages.

3.5. Chapter Summary and Future Directions

In this chapter, we take the first step towards identifying and analyzing empathetic messages in online health communities. We proposed a model based on a combination of Convolutional Neural Networks and Long Short Term Memory, called ConvLST, for empathetic messages identification. We showed that the proposed ConvLST model outperforms several baselines. Moreover, we experimentally showed that empathetic messages are one of the main reasons for changing participants' feelings (from negative to positive) in OHCs.

In future, it would be interesting to investigate the relation between the number of originators' posts across different discussion boards and the probability of changes in his/her feelings.

CHAPTER 4

ANALYZING INFORMATIONAL MESSAGES IN OHCS

[1] Online health communities have become a medium for patients to share their personal experiences and interact with peers on topics related to a disease, medication, side effects, and therapeutic processes. Analyzing informational posts in these communities can provide an insightful view about the dominant health issues and can help patients find the information that they need easier, and ultimately, cope better with a disease.

In this chapter, we propose a computational model that mines user content in online health communities to detect positive experiences, advices and suggestions on health improvement as well as negative impacts or side effects that cause suffering throughout fighting with a disease. Specifically, we combine high-level, abstract features extracted from a convolutional neural network with lexicon-based features and features extracted from a long short term memory network to capture the semantics in the data. We show that our model, with and without lexicon-based features, outperforms strong baselines. We also compared our model with several deep neural network and traditional machine learning approaches and showed that our model achieved,by far, the best results.

4.1. Introduction

Traditionally, medical doctors and care providers have been the main source of information for patients who suffer from chronic or life-threatening diseases. However, with the advent of the Internet and the creation of many online health communities (OHCs), e.g., Everyday Health, Cancer Survivors' Network, and WebMD, patients use these health communities increasingly as an integral source for finding health-related information [23, 34]. OHCs provide an environment for patients, their family members and friends to interact with other participants and share experiences and information (e.g., recommendations and feedback) on issues related to prescribed medicines, side effects, therapeutic processes, mental health, and feelings. Table 4.1 shows examples of posts

[1]This entire chapter is reproduced from Khanpour, H., Caragea, C. [2018], Fine-Grained Information Identification in Health Related Posts, in 'The 41st International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)', currently under review.

Exp. 1: I am on Sertraline, which is generic zoloft and I truly believe it has helped me. I have been on it since I was initially diagnosed. I took it all thru chemo and I am still on it. Doctors also say it helps with hot flashes. I don't know about that since I still get them. But at least I am not depressed. So that is good. **Exp. 2 :**The antidepressants need time to build up in your system. Several weeks or a month to get relief and they seem to work for 6 months to a year. The anti anxieties (xanax, valuim, ativan) are for instant relief of stress and anxiety. They are habit forming. I have been on Xanax (alprazolam) for more than 10 years and my dosage has increased since my DX in June of this year. I am trying to back off on them some. After going through Cancer DX, Surgery, Chemo i think we all need some kind of mental health help.
Exp. 3: I took Anzamet... one pill prior to the infusions and one each day for 3 days following treatments. The only real problem I developed with it, and it lasted till I finished treatment, was an aversion to drinking water! Plain water began to just taste terrible to me. **Exp. 4:** I went through 6 months of chemo and then within 2 weeks of radiation I developed upper right back pain. It feels like a pinched nerve. I've been told it was due to the positioning during radiation, or scar tissue or the fact that I now carry everything on my left and I might be not standing as straight.

TABLE 4.1. Examples of OHC informational messages.

that contain health-related information shared among patients in an online cancer community. This information is very unique and is often not available elsewhere, e.g., referring to the medication Sertraline, a patient writes: *Doctors also say it helps with hot flashes. I don't know about that since I still get them* (see Example 1 in the table).

Several research studies showed that using OHCs to obtain information from people who went through the same or similar experiences (either by direct interactions or sifting through the online posts) adds substantial value to it as it brings better feelings and fewer mortality odds to patients [48, 82]. Thus, the large and growing amounts of user-generated content in OHCs need to be accurately classified for a variety of applications, e.g., designing smart information retrieval systems for content recommendation, and improving data organization for fast retrieval. Recent computational studies in OHCs started to investigate the high level identification of informational posts [10, 114], however, with no emphasis on the unique challenges associated with the detection of the information type, e.g., therapeutic procedures vs. side effects. A deep understanding of the text and the writer's intention is required in order to correctly extract the types of information present in messages posted in OHCs. Example 1 and Example 2 in Table 4.1 refers to therapeutic procedure, whereas Example 3 and Example 4 refers to side effects through various medication (Sertraline and Anzamet, respectively).

In this chapter, we propose to analyze messages in OHCs to extract the information type that they contain, i.e., *therapeutic procedures* (any medical treatment, activity, or behavior that have a positive impact on patients' health, precisely, can help prevent, cure or improve a patient's condition) and *side effects* (any medical treatment, activity, or behavior that have a negative impact on patients' health, precisely, a secondary, often undesirable effect of a drug or medical treatment). To achieve this, we design a computational model that is able to exploit the semantic information from text, and coherently combines high-level (abstract) features with surface-level and lexicon-based features. Our contributions are as follows:

(1) We propose to extract fine-grained information types from messages posted in online health communities. Identifying information types provides doctors, health practitioners and OHCs' moderators with an insightful view of patients' physical status during various treatments. In addition, it can provide new diagnosed patients with information about what they should expect throughout their treatments and help them in making informed decisions about their disease more effectively [98, 99]. To our knowledge, we are the first to address fine-grained information type extraction in OHCs.

(2) We design and explore a computational model that can identify messages belonging to *therapeutic procedures* and *side effects* with high accuracy. Our model is a hybrid neural network combined with lexicon-based features, which we call HNNL. HNNL combines the output of a Convolutional Neural Network (CNN) with the output of a Long Short-

Term Memory (LSTM) network and lexicon-based features, which are all fed into a fully connected network with SoftMax layer.

(3) We show empirically that HNNL significantly outperforms strong baselines and prior works; moreover, we show that the proposed model continues to perform well even in the absence of lexicon-based features.

4.2. Data Collection and Annotation

Since there is no available dataset for analyzing messages that contain fine-grained informational content in OHCs (i.e., therapeutic procedures and side effects), we constructed a benchmark dataset to evaluate our model. We randomly selected 225 comments from 21 discussion threads in the Lung cancer discussion board and 120 comments from 11 discussion threads in the prostate cancer discussion board of the ACS Cancer Survivors' Network (CSN). Consistent with Biyani et al. [10] who performed sentence level classification since long messages often comprise of different topics, we performed our data annotation at sentence level. Following the same study [10], we selected sentences with length greater than four words to exclude appreciative and appraisal messages. We obtained $1,797$ sentences, which were integrated with the $1,066$ sentences extracted from the breast cancer discussion board in CSN that were provided by Biyani et al. 2014, with $2,863$ overall sentences.

The purpose of the annotation was to label the $2,863$ sentences as belonging to *therapeutic procedures*, i.e., containing information about any medical treatment, activity, or behavior that have a positive impact on patients' health; *side effects*, i.e., containing information about any medical treatment, activity, or behavior that have a negative impact on patients' health; and *other*, which includes sentences that do not belong to any of the above two categories. Two annotators (graduate students) contributed to the task. They were trained by the first author of this study through several trial tasks during 10-days period. The annotation task was conducted in an iterative fashion following prior studies and guidelines [31, 35, 97]. In each round, 200 messages were assigned and annotators discussed disagreements with researchers; 100% inter-annotator agreement (IAA) was achieved after each round of discussions. We used Cohen's kappa for measuring IAA. After four initial rounds of annotations, the remaining data ($2,063$ messages) were assigned to the annotators where they achieved 90% IAA. The last round of the assigned data was adjudicated by one of the authors.

Table 4.2 provides the distribution of messages in each category. As can be seen, *therapeutic procedures* has significantly more sentences than *side effects*. This shows that patients tend to share more of their success stories and positive aspects of their therapy rather than sharing negative impacts or side-effects. The category *other* has the largest number of sentences. A large fraction of these sentences contain emotional support such as empathy, and encouragements.

Category	**#sentences**	**percentage (%)**
therapeutic procedures	942	32.9
side effects	385	13.4
other	1536	53.7

TABLE 4.2. Statistics from the data collection.

4.3. Model

In this section, we describe our proposed computational model which can be embedded in OHCs' search engines to retrieve fine-grained messages belonging to *therapeutic procedures* or *side effects*. Given a sentence of n words, we apply Convolutional Neural Networks (CNN) and Long Short Term Memory (LSTM) networks concurrently. The CNN extracts high-level (abstract) features that capture the semantic part of the text [62] whereas the LSTM captures sequential information from each sentence.

Our CNN architecture consists of one convolution layer followed by a max pooling layer. The input data layer is fed with word vectors of length k, where $x_i \in \mathbb{R}^k$ is the k-dimensional word vector corresponding to the i-th word in the sentence. Thus, the input sentence of length n is represented as

$$x_{1:n} = x_1 \oplus x_2 \oplus .. \oplus x_i \oplus ... \oplus x_n \tag{11}$$

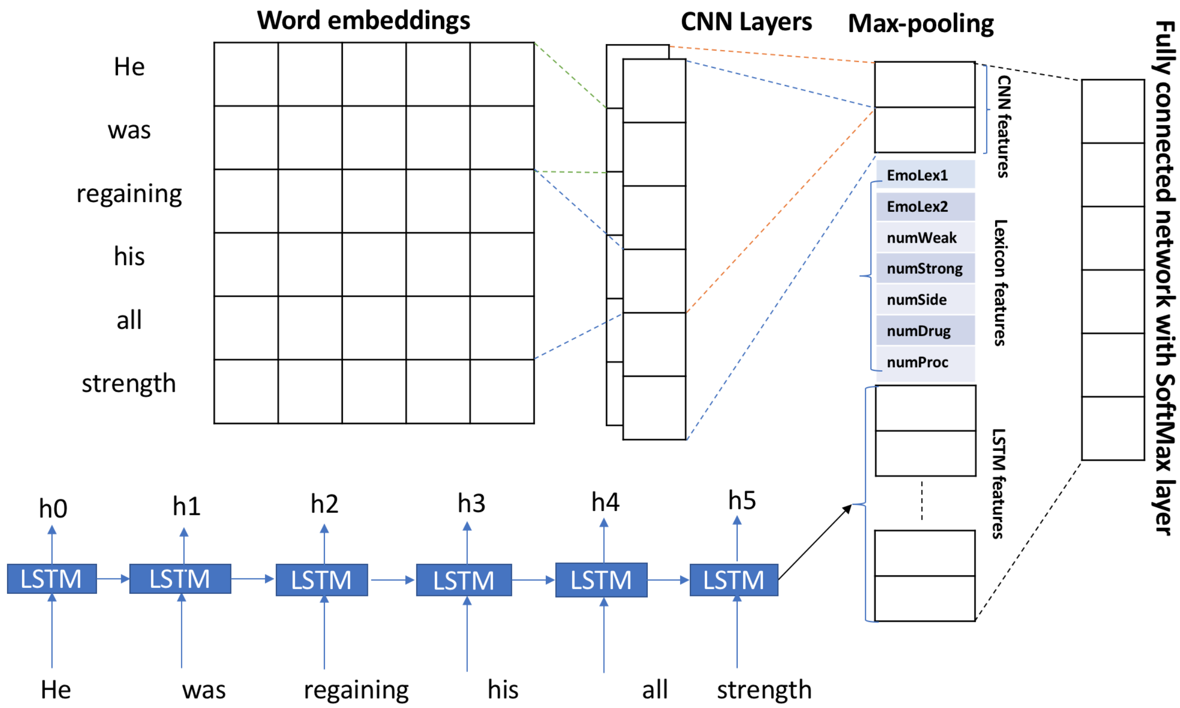

FIGURE 4.1. The architecture of our proposed hybrid neural network with lexicon-based features (HNNL).

where $\oplus$ is the concatenation operator. First, l filters for each region size (rs) are applied to the sequence of tokens in each sentence (e.g., 16 filters for 2 and 3 region sizes). The feature map M belongs to $\mathbb{R}^{l \times rs}$. The features m_j, with $j = 1, \cdots, T$ (where T is the number of extracted features), are defined as follows:

(12) $$m_j = Relu(M[x_{j-(rs/2)+1}, \cdots, x_j, \cdots, x_{j+(rs/2)}])$$

where $Relu$ is the rectified linear unit activation function. This process is iteratively done for each time step (corresponding to each word) of the input sentence that ends up with $M = (m_1, m_2, \cdots, m_T)$ sequence. Second, max pooling is applied to M, which results in $M' = (m_1{'}, m'_2, ..., m'_{T/(pooling-size)})$. M' is the output of the CNN that contains high-level, abstract features.

The LSTM unit consists of sub-unit-inputs (i_t), output (o_t), forget gates (f_t) and memory cell (c_t).

$$i_t = \sigma\left(W^{(i)}x_t + U^{(i)}h_{t-1} + b^{(i)}\right) \tag{13}$$

$$f_t = \sigma\left(W^{(f)}x_t + U^{(f)}h_{t-1} + b^{(f)}\right) \tag{14}$$

$$o_t = \sigma\left(W^{(o)}x_t + U^{(o)}h_{t-1} + b^{(o)}\right) \tag{15}$$

$$u_t = \tanh\left(W^{(u)}x_t + U^{(u)}h_{t-1} + b^{(u)}\right) \tag{16}$$

LSTM unit at time t computes the memory cell:

$$u_t = \tanh(Wx_t + Uh_{t-1} + b) \tag{17}$$

$$c_t = i_t \bigodot u_t + f_{(t)} \bigodot c_{t-1} \tag{18}$$

and then computes the output, or activation:

$$h_t = o_t \bigodot \tanh(c_t) \tag{19}$$

Here, $x \in \mathbb{R}^{n \times k}$ is the input and $W \in \mathbb{R}^{n \times k}$, $U \in \mathbb{R}^{n \times n}$, and $b \in \mathbb{R}^n$ are parameters of an affine transformation. The resulting sequence of the layers is $h_1, h_2, \cdots, h_n$.

Last, we combine the features extracted by CNN and LSTM networks with lexicon-based features in a hybrid model. Our proposed model, HNNL, is shown graphically in Figure 4.1. As can be seen from the figure, we use a combination of CNN and LSTM models, where the final feature vectors from CNN augmented with lexicon-based features and the last feature h_n extracted by LSTM are fed into a fully connected network with a SoftMax layer.

Lexicon-based Features: In this work, we used seven lexicons. The first five lexicons come from Biyani et al. 2014. These lexicons are: weak subjective words (**numWeak**), strong subjective words (**numStrong**) cancer drugs (**numDrug**), side-effects (**numSide**), and therapeutic procedures (**numProc**). The sixth and seventh lexicons come from emotion detection research. Our motivation for the integration of these two emotion lexicons is that a large fraction of sentences in the category *other* are emotional in nature, where people emotionally support one another. These lexicons are: **EmoLex1** [101] and **EmoLex2** [80]. We use frequencies of lexicon words to construct the lexicon-based features.

4.4. Experiments and Results

Next, we describe the evaluation of HNNL using binary tasks. Specifically, we trained our models in the two-class setting by binarizing the datasets: *therapeutic procedures* vs. *non-therapeutic procedures* (and *side effects* vs. *non-side effects*). In all experiments, we used word embeddings as input to the neural networks, which were generated with the W2vector module in Gensim [92] on the data from all discussion boards of CSN between 2000 and 2012. The results (weighted average Precision, Recall and F1-score) are reported in 10-fold cross validation experiments.

Hyper-parameter setting: We optimized hyper-parameter values of our HNNL model as well as all the other neural network models (used for comparison) by performing a grid search on a development set, which consists of 20% of instances removed from the training set in each iteration of 10-fold cross-validation. Table 4.3 shows the best hyper-parameter values for all neural network models.

Performance of HNNL in an ablation experiment.

Since our HNNL model is a hybrid neural network model with several components, first, we evaluate its performance in an ablation experiment to understand the contribution of each component in the model performance. Specifically, we compare HNNL with HNN (a model that has the same architecture as HNNL, but does not use any external lexicon), CNN, LSTM, and support vector machines (SVM) with the (concatenated) features from the seven lexicons (described

HNNL: LR= 0.1, Decay rate=0.6, Dropout=0.8, Layer=1, Max pooling, FRS=(2,3), NF=16
HNN: LR= 0.1, Decay rate=0.6, Dropout=0.8, Layer=1, Max pooling, FRS=(2,3), NF=16
LSTM: LR= 0.001, L2reg=1$E-5$, Decay rate=0.7, Layer=1, Max pooling
CNN: LR= 0.1, Decay rate=0.5, Dropout=0.6, Layer=2, Max pooling, FRS=(2,3,4), NF=16
ConvLSTM: LR= 0.1, Decay rate=0.7, Dropout=0.6, Layer=2, Max pooling, FRS=(2,3), NF=16
ConvLexLSTM: Decay rate=0.8, Dropout=0.5, Layer=1, Max pooling, FRS=(2,3), NF=16

TABLE 4.3. Hyper-parameter settings for all models.

Method	TP			SE		
	Pr	**Re**	**F1**	**Pr**	**Re**	**F1**
HNNL	**82.9**	**87.1**	**84.9**	**79.8**	**81.3**	**80.5**
HNN	80.3	84.6	82.3	77.5	79.3	78.4
LSTM	75.8	78.2	76.9	71.2	73.0	72.0
CNN	70.0	72.3	71.1	68.0	69.7	68.8
Seven-Lexicon	69.2	65.0	67.0	65.2	67.1	66.1
C-ConvLSTM	78.1	75.6	76.8	76.6	71.0	73.6
LibShortText toolkit6	69.9	72.4	71.1	68.5	69.6	69.0
Tf-Idf	62.7	63.8	63.2	60.1	63.0	61.5
ConvLexLSTM	79.5	82.9	81.1	76.4	77.7	77.0
ConvLSTM	78.6	79.9	79.2	73.7	77.0	75.3

TABLE 4.4. Classification results of HNNL vs. other models. TP denotes *therapeutic procedures*, and SE denotes *side effects*.

above).

Table 4.4 shows the results of these comparisons (first block of results). As can be seen, HNNL achieves the best results consistently throughout all experiments in terms of all compared measures. This ablation experiment confirms that all components in our model positively contribute to the final results. For example, eliminating the seven lexicon features from HNNL, which yields HNN, results in a drop in F1-score by 2.6% on *therapeutic procedures* and by 2.1% on *side effects* classification results. Still, HNN is the second performing model in terms of F1-score. These results show that our model can be successfully applied in a health domain even in the absence of health lexicons, which are often expensive to obtain and require domain experts to design them. Not surprisingly, the SVM with the seven lexicon-based features (denoted as Seven-Lexicon) performs the worst among the compared models, suggesting that obtaining the semantic information from text improves models' performance.

Baseline Comparisons.

Second, we compare HNNL with several baselines and prior works: (1) C-ConvLSTM (i.e., a character-level CNN-LSTM) by Kim et al. [60] in which the output of CNN is input for LSTM; (2) LibShortText toolkit6 that uses SVM with part of speech tags and other syntactic and semantic features such as frame-semantics, dependency triples, and (3) an SVM with *tf-idf* features. The LibShortText toolkit6 has been shown to have a very good performance on classifying short-texts [111]. Table 4.4 shows the results of this comparison in the second block of results. As can be seen, HNNL and HNN outperform all three baselines, and more importantly, they outperform the C-ConvLSTM, which represents a different (i.e., serial) combination of CNN and LSTM networks, however, at character level. Interestingly, how would the HNNL model that uses CNN and LSTM in parallel compare with a (serial) combination of CNN and LSTM at word level? To understand this, we designed a model called ConvLSTM, i.e., a word-level CNN-LSTM that uses word embeddings trained on CSN data instead of character-level CNN-LSTM as in C-ConvLSTM. We also extended ConvLSTM to include the seven lexicons into the ConvLexLSTM model. More precisely, in ConvLexLSTM, the CNN features augmented with lexicon-based features are fed as input to LSTM. The results of these two baselines are shown in the last block of Table 4.4. As can be seen from the table, the performance of word-level ConvLSTM is higher than character-level

C-ConvLSTM. Adding the lexicon features to ConvLSTM yields even higher performance, but not as high as that of HNNL. This result confirms our belief that preserving the sequential information in sentences added to the CNN features yields improvement in performance over models which do not include sequential information.

It is also worth mentioning that all deep neural networks that capture semantics from the data perform better that the traditional models. The lexicon-based features act as a complement (for the high-level semantic features) by finding exact words in the text to generate proper features in HNNL.

4.5. Related Work

In computational studies, messages in OHCs have been analyzed from the standpoint of social support. Among the variety of types of social support, i.e., emotional, informational, instrumental, and appraisal support [30, 39], the two types of support, emotional [34] and informational, [11] form the two principle functions that shape the majority of messages in OHCs. Thus far, most computational studies in OHCs are dedicated to analyzing and identifying messages that contain these two types of support. For example, Wang et al. [114] used a linear regression model to identify emotional and informational support in messages from a cancer forum and studied the relationship of these support types on user engagement with the health community. Their feature set for the regression model includes: LIWC features, POS tags, message length, subjectivity intensity, and Latent Dirichlet Allocation based topical features.

Biyani et al. [10] used classification models (e.g., Naïve Bayes and Logistic Regression) to classify messages that contain emotional or informational support from posts in a breast cancer discussion board of a cancer survivors' network. The authors used unigrams, POS tags, structural linguistic patterns, and five lexicons that contain strong and weak subjective words, cancer drugs, side-effects, and cancer procedures, and showed that features drawn from lexicons have the highest impact on the results. Their assumption was that in the informational messages, patients talk about their personal experiences with the disease that most likely include a word from one of the five lexicons. On a breast cancer dataset constructed from the Cancer Survivors' Network (CSN) of the American Cancer Society (ACS), the authors showed that their classifier can identify emotional and informational messages with an F1-score of 0.88 and 0.77, respectively. Similar to Wang et al. [114], Wang et al. [113] studied the correlation between social support and user engagement, but instead of using a regression model, the authors used traditional machine learning classifiers such as Naïve Bayes, Logistic Regression, Support Vector Machines, Random Forest to classify OHCs' messages based on the intention of the participant when writing a message (i.e., companionship, seeking information, seeking emotion, providing information, and providing emotion). The authors used a combination of features from Wang et al. [114] coupled with lexicon-based features used in Biyani et al. [10].

In contrast to the above works that have been devoted to high level analyses of emotional and informational messages, we focus on the unique challenges associated with fine-grained detection of informational messages, i.e., messages belonging to the categories *therapeutic procedures* and *side effects*, that have the potential to improve patients' competence and knowledge in dealing with health care problems and will empower them to become better prepared and take control of their life in better ways.

4.6. Chapter Summary and Future Directions

In this chapter, we proposed a computational model for classifying fine-grained informational messages in OHCs. Our proposed model, HNNL, combines the strengths of CNNs, LSTMs, and lexicon-based approaches to capture hidden semantics in OHCs' messages. We show that our proposed model, with or without lexicon-based features, which are often expensive to obtain or maintain in a health domain, provides a better computational model for classifying informational messages based on their content compared with strong baselines, including other types of deep neural networks. In future, it would be interesting to study the performance of our models on data from different health communities, e.g., related to weight loss.

CHAPTER 5

CONCLUSION

In this chapter, we summarize the the contributions of this work and present future direction in our research.

5.1. Dissertation Summary

Recent advances in online applications and computer technologies have persuaded people to explore web pages or run some queries to obtain required information. Studies show that the initial source for patients to obtain information about their disease has changed gradually such that Internet resources has become the first source of obtaining medical and health-related information. Recently, by emerging Web 2.0 applications on the Internet, many online applications and forums have been developed to serve patients with their needs. Online health communities (OHCs) are a type of forum that provides a variety of environments for patients and their family and friends. OHc participants can communicate with each other and exchange their knowledge and experiences on disease, medical information, side-effects, etc.. Emotional and informational support are considered as the two important benefits of using OHCs. Many studies investigated the impact of using OHCs and postulated that using OHCs regularly decreases the chance of mortality and improves patients' mental health [48, 61]. As a result of their benefits, OHCs are a popular place for patients to refer to, especially patients with a severe disease.

Studies show that patients use OHCs to receive emotional and informational support; the main reasons for developing OHCs are to present valid and high-quality information and to understand the mechanism of social support functioning on patients [12, 109].

Given the purpose of OHC moderators for developing OHCs' applications and the purpose of patients for using OHCs, there is no facility, feature, or sub-application in OHCs to satisfy patient and moderator goals. OHCs are only equipped with a primary search engine that is a keyword-based search tool. In other words, if a patients wants to obtain information about a side-effect, he/she needs to browse many threads in the hope that he/she can find several related comments. In the same way, OHC moderators cannot browse all information which is exchanged among patients to validate their accuracy. Thus, it is critical for OHCs to be equipped with some tools which are supported by several sophisticated computational models that provide moderators and patients with messages that they need for making decisions or predictions. For example, a model could only provide patients with messages that explain side-effects and therapeutic processes, or only extract emotional messages with sad content so that moderators could identify depressed patients who post many messages with sad content.

Through this study, we present multiple computational models to alleviate the problem of OHCs in providing specific types of messages in response to the specific moderator and patient needs. Specifically, we focused on classification based on the following categories:

- Emotional support, which presents OHCs moderators, psychologists, and sociologists with insightful views on the emotional state's of individuals and groups.
- Empathy, which gives an insightful view and useful means for OHC moderators, psychologists, and sociologists to understand the mechanism of social support in OHCs.
- Informational support, which provides patients with an efficient and effective tool for accessing the best-fit messages throughout huge amount of patient posts. This classifier also provides an insightful view to OHC moderators, health-practitioners, nurses, and doctors about the current discussion under the topics of side-effects and therapeutic processes, giving them an opportunity to monitor and validate the exchange of information in OHCs.

The following is a summary of this dissertation:

- **Emotional Messages Identification and Emotion Type Detection in OHC Messages:** In this work, we extended previous datasets provided by prior studies on emotional message identification. We also built two new datasets for the task of emotion type detection. We showed that men are more emotional when they post messages in forums where the majority of members are female. We presented a classifying model for emotional vs. non-emotional messages and emotion type detection, showing that this model performs the classification task with high accuracy even without using any contextual information. Two comparison studies were performed that showed us our computational models outperformed previous studies. We ran an experimental study to investigate patients' emotional health around important holidays in the US. We collected and analyzed an OHC's (i.e., Cancer Survivors Network) participant posts during a 10-year period. We showed that on a typical day, the percentage of joy and sad emotions are similar, whereas on Christmas and Thanksgiving the percentages show more joyful spirits, possibly due to family gatherings and other social events around these holidays, in which people feel supported and hence feel better. We also collected influential users' posts and found that their posts in OHCs have considerably more joy content than sad content.

- **Identifying Empathetic Messages in Online Health Communities (OHCs):** we investigated the role of empathetic messages in social media and specifically in OHCs by proposing a computational model for identifying empathetic messages in OHCs. Since there was no dataset for empathy in OHCs, we built a dataset using messages posted in two forums (i.e., breast cancer and lung cancer) of the Cancer Survivors Network (CSN). We developed a computational model for classifying messages in OHCs. Our model creates high-level and abstract features by using the Convolutional Neural Network (CNN); each of the extracted features is fed into a single Long-short Term Memory (LSTM) cell that classifies messages based on the extracted features using Softmax. We also ran an experiment using the whole messages posted in CSN from June 2000 until June 2012. We applied our model to the extracted data and found that empathetic messages are the reason for 39% of positive shifts in patient health.
- **Analyzing Informational Messages in Online Health Communities:** In this work, we addressed the problem of analyzing informational messages by classifying them based on their content. Similar to empathetic message analysis, this work also presented a detailed analysis of informational messages for the first time in the literature. We built a dataset from messages that were randomly extracted from breast cancer and lung cancer forums of the Cancer Survivors Network (CSN). We annotated 2107 messages to highlight messages that contained side-effects and therapeutic processes. In this study, specifically, we addressed the following problems:

- Identifying messages in OHCs which contain any mentions of therapeutic processes–including any medical treatment, activity, or behavior that have a positive impact on patient health–can help prevent, cure, or improve a patient's condition.
- Identifying messages in OHCs which contain any mentions of side-effects–including any medical treatment, activity, or behavior that have a negative impact on patients' health. Precisely, a secondary, often undesirable effect of a drug or medical treatment.

We designed a computational model which exploits three types of features from three different sources for the classification task. The first source of features is CNN-based which generates high-level, abstract features that address semantic part of the messages. The second source of features is constructed by using surface-level analysis and lexicon-based features; theses lexicons contain emotional words, weak and strong subjective words, and lists of side-effects, drugs, and therapeutic procedures. Finally, the third source of features are provided by an LSTM network that captures sequential information in messages. We showed that our model outperformed strong baselines.

5.2. Summary of Contributions

This section presents the contributions of our works in this dissertation and outlines some of the future directions of research in this domain:

- **Emotional Messages Identification and Emotion Type Detection in OHC Messages:**
  - We proposed to detect emotion types in messages posted in online health communities. Identifying emotion types in patients' messages improves the capability of OHC moderators, caregivers, and doctors to provide high-quality services to OHC users or patients. To our knowledge, we are the first to address emotion type detection in OHCs.
  - We proposed a computational model, called ConvLexLSTM, for emotional messages identification and emotion detection in OHCs messages. Our model combines the output of a Convolutional Neural Network (CNN) with lexicon-based features, which are all fed into a Long Short-Term Memory (LSTM) network that generates the final output using softmax function. We empirically showed that ConvLexLSTM sig-

nificantly outperforms strong baselines and the state-of-the-art models presented by prior studies. Moreover, we showed that the proposed model continues to perform well even in the absence of lexicon features (i.e., ConvLSTM).

  - Finally, we applied ConvLexLSTM to a large scale experiment to study the correlation between US holidays and users' emotional states, which can help design smarter approaches to improve patients' moods.
  - As part of our contributions, we constructed two datasets for emotion type detection and extended the dataset built for emotional messages identification in OHCs. To our knowledge, our datasets for emotion type detection are the first constructed for this task.

- **Identifying Empathetic Messages in Online Health Communities (OHCs):** We studied empathetic messages in OHC messages. The results of this study is specially useful for OHCs moderators, and researchers in the field of psychology and sociology. Our contributions in this study are:
  - We built a dataset for empathetic messages identification in OHCs. This was the first study which provided a dataset for identifying empathetic messages in a health-related social media.
  - We proposed an automatic model for identifying empathetic messages in OHCs. To the best of our knowledge, this is the first work that proposes a computational model for detecting empathy in social networks. Our model used generated features from CNN network and feeding them into an LSTM network.
  - We experimentally validated our empathy identification model on a manually annotated dataset specifically created for this task from a Cancer Survivors' Network.
  - We also showed that generally empathetic messages in comments of OHCs improve participants feelings from negative to positive.
- **Analyzing Informational Messages in Online Health Communities:**
  - We proposed to extract fine-grained information types from messages posted in online health communities. Identifying information types provides doctors, health prac-

titioners and OHCs' moderators with an insightful view of patients' physical status during various treatments. In addition, it can provide new diagnosed patients with information about what they should expect throughout their treatments and help them in making informed decisions about their disease more effectively [98, 99]. To our knowledge, we were the first to address fine-grained information type extraction in OHCs. We designed and explored a computational model that can identify messages belonging to *therapeutic procedures* and *side effects* with high accuracy. Our model is a hybrid neural network combined with lexicon-based features, which we call HNNL. HNNL combines the output of a Convolutional Neural Network (CNN) with the output of a Long Short-Term Memory (LSTM) network and lexicon-based features, which are all fed into a fully connected network with SoftMax layer.

– We showed empirically that HNNL significantly outperformed strong baselines and prior works; moreover, we showed that the proposed model continues to perform well even in the absence of lexicon-based features.

5.3. Future Directions

As we mentioned before, the main purpose of creating OHCs is to provide patients with validated information about their disease and understanding the mechanism of social support functions on patients and the main purpose of patients in using OHCs is to receive emotional and informational support. In this study, we take initial steps in improving current primitive structure of OHCs to satisfy the main goals of OHC patients in using OHCs and also the main goals of OHC creators in building OHCs. As immediate next steps we plan to:

- Develop a computational model for identifying influential users in OHCs. In this study we proposed two models which can be used for future studies in this direction (i.e., emotion type detection and empathetic messages identification). Finding influential users can help moderators with validating or correcting exchanging information between patients.
- Design compositional models for detecting and tracking depressed patients. In this study we proposed a computational model for detecting emotion type in patients' post. However, detecting depressed patients needs more factors to be included such as patterns of

participations in OHCs, the degree of their engagement in the discussion, the degree of optimism and pessimism in his/her view and so on.

- Create a big dataset that contain heterogeneous data from different forums in OHCs. Building a big dataset allows us to build more convoluted computational models with more accurate results. Also, working on heterogeneous dataset allows us to investigate and compare individuals' characteristics in different ways; for example, studying the difference between genders in posting empathetic messages in different forums–gender studies will help OHC moderators and researchers in studying social support mechanism in OHCs.
- Design some computational models to extract side-effects of a specific disease. Currently our models extract messages in OHCs that contain side-effects and therapeutic process. However, we can go deeper and find the exact mention of the side-effects and therapeutic process in messages. These models have the potentials to help moderators in monitoring information and help patients in accessing their information efficiently.
- We intend to create lexicons of words and abbreviations that convey a sentiment in OHCs. Since the main topic of discussions in OHCs are about health and disease, the words and abbreviations that are used in OHCs are likely to have different meaning or sentiments. For example, the word *hot flashes* convey a quite negative sentiment in OHCs, nevertheless in general language that might not convey a specific sentiment.

REFERENCES

[1] Agrawal, A. and An, A. [2012], Unsupervised emotion detection from text using semantic and syntactic relations, *in* 'Proceedings of the The 2012 WI-IAT Volume 01', pp. 346–353.

[2] Bahdanau, D., Cho, K. and Bengio, Y. [2014], 'Neural machine translation by jointly learning to align and translate', *arXiv preprint arXiv:1409.0473* .

[3] Ballesteros, M., Goldberg, Y., Dyer, C. and Smith, N. A. [2016], 'Training with exploration improves a greedy stack-lstm parser', *arXiv preprint arXiv:1603.03793* .

[4] Bao, S., Xu, S., Zhang, L., Yan, R., Su, Z., Han, D. and Yu, Y. [2009], Joint emotion-topic modeling for social affective text mining, *in* 'Data Mining, 2009. ICDM'09.', pp. 699–704.

[5] Barrera Jr, M. [2000], Social support research in community psychology, *in* 'Handbook of community psychology', Springer, pp. 215–245.

[6] Batson, C. D. [2009], 'These things called empathy: eight related but distinct phenomena.'.

[7] Bender, J. L., Katz, J., Ferris, L. E. and Jadad, A. R. [2013], 'What is the role of online support from the perspective of facilitators of face-to-face support groups? a multi-method study of the use of breast cancer online communities', *Patient education and counseling* **93**(3), 472–479.

[8] Bengio, Y., Simard, P. and Frasconi, P. [1994], 'Learning long-term dependencies with gradient descent is difficult', *IEEE transactions on neural networks* **5**(2), 157–166.

[9] Bishop, C. M. [2006], *Pattern Recognition and Machine Learning (Information Science and Statistics)*, Springer-Verlag New York, Inc.

[10] Biyani, P., Caragea, C., Mitra, P. and Yen, J. [2014], Identifying emotional and informational support in online health communities., *in* 'COLING', pp. 827–836.

[11] Boon, H. S., Olatunde, F. and Zick, S. M. [2007], 'Trends in complementary/alternative medicine use by breast cancer survivors: comparing survey data from 1998 and 2005', *BMC women's health* **7**(1), 4.

[12] Bouma, G., Admiraal, J. M., de Vries, E. G., Schröder, C. P., Walenkamp, A. M. and Reyn-

ers, A. K. [2015], 'Internet-based support programs to alleviate psychosocial and physical symptoms in cancer patients: a literature analysis', *Critical reviews in oncology/hematology* **95**(1), 26–37.

[13] Boyd, R. L., Wilson, S. R., Pennebaker, J. W., Kosinski, M., Stillwell, D. J. and Mihalcea, R. [2015], Values in words: Using language to evaluate and understand personal values, *in* 'Ninth International AAAI Conference on Web and Social Media'.

[14] Brownstein, C. A., Brownstein, J. S., Williams III, D. S., Wicks, P. and Heywood, J. A. [2009], 'The power of social networking in medicine', *Nature biotechnology* **27**(10), 888.

[15] Castleton, K., Fong, T., Wang-Gillam, A., Waqar, M. A., Jeffe, D. B., Kehlenbrink, L., Gao, F. and Govindan, R. [2011], 'A survey of internet utilization among patients with cancer', *Supportive Care in Cancer* **19**(8), 1183–1190.

[16] Centola, D. [2010], 'The spread of behavior in an online social network experiment', *science* **329**(5996), 1194–1197.

[17] Chiu, J. P. and Nichols, E. [2015], 'Named entity recognition with bidirectional lstm-cnns', *arXiv preprint arXiv:1511.08308* .

[18] Christakis, N. A. and Fowler, J. H. [2007], 'The spread of obesity in a large social network over 32 years', *New England journal of medicine* **357**(4), 370–379.

[19] Cohen, S., Underwood, L. G. and Gottlieb, B. H. [2000], *Social support measurement and intervention: A guide for health and social scientists*, Oxford University Press.

[20] Collins, F. M. [2014], 'The relationship between social media and empathy'.

[21] Collobert, R. and Weston, J. [2008], A unified architecture for natural language processing: Deep neural networks with multitask learning, *in* 'Proceedings of the 25th international conference on Machine learning', ACM, pp. 160–167.

[22] Conneau, A., Schwenk, H., Barrault, L. and Lecun, Y. [2016], 'Very deep convolutional networks for natural language processing', *arXiv preprint arXiv:1606.01781* .

[23] Coulson, N. S., Buchanan, H. and Aubeeluck, A. [2007], 'Social support in cyberspace: a content analysis of communication within a huntington's disease online support group', *Patient education and counseling* **68**(2), 173–178.

[24] Coursaris, C. K. and Liu, M. [2009], 'An analysis of social support exchanges in online hiv/aids self-help groups', *Computers in Human Behavior* **25**(4), 911–918.

[25] Das, A. and Faxvaag, A. [2014], 'What influences patient participation in an online forum for weight loss surgery? a qualitative case study', *Interactive journal of medical research* **3**(1).

[26] Dauphin, Y. N., Fan, A., Auli, M. and Grangier, D. [2016], 'Language modeling with gated convolutional networks', *arXiv preprint arXiv:1612.08083* .

[27] Davis, M., Tiedens, L. and Leach, C. [2004], 'Negotiating the border between self and other', *The social life of emotions* pp. 19–42.

[28] Decety, J. and Jackson, P. L. [2004], 'The functional architecture of human empathy', *Behavioral and cognitive neuroscience reviews* **3**(2), 71–100.

[29] Del Rey, R., Lazuras, L., Casas, J. A., Barkoukis, V., Ortega-Ruiz, R. and Tsorbatzoudis, H. [2016], 'Does empathy predict (cyber) bullying perpetration, and how do age, gender and nationality affect this relationship?', *Learning and Individual Differences* **45**, 275–281.

[30] Dennis, C.-L. [2003], 'Peer support within a health care context: a concept analysis', *International journal of nursing studies* **40**(3), 321–332.

[31] DMello, S. K. [2016], 'On the influence of an iterative affect annotation approach on inter-observer and self-observer reliability', *IEEE Transactions on Affective Computing* **7**(2), 136–149.

[32] Ekman, P. [1992], 'An argument for basic emotions', *Cognition & emotion* **6**(3-4), 169–200.

[33] Elhadad, N., Zhang, S., Driscoll, P. and Brody, S. [2014], Characterizing the sublanguage of online breast cancer forums for medications, symptoms, and emotions, *in* 'AMIA Annual Symposium Proceedings', Vol. 2014, American Medical Informatics Association, p. 516.

[34] Eysenbach, G., Powell, J., Englesakis, M., Rizo, C. and Stern, A. [2004], 'Health related virtual communities and electronic support groups: systematic review of the effects of online peer to peer interactions', *Bmj* **328**(7449), 1166.

[35] Fort, K. et al. [2016], *Collaborative Annotation for Reliable Natural Language Processing: Technical and Sociological Aspects*, Wiley Online Library.

[36] Fowler, J. and Christakis, N. [2008], 'Dynamic spread of happiness in a large social network:

longitudinal analysis over 20 years in the framingham heart study. bmj, 337 (dec04 2), a2338–a2338'.

[37] Fox, S. [2013], 'Pew research centers internet and american life project. report health online 2013, california healthcare foundation'.

[38] Friedenreich, C. M., Neilson, H. K., Farris, M. S. and Courneya, K. S. [2016], 'Physical activity and cancer outcomes: a precision medicine approach', *Clinical Cancer Research* **22**(19), 4766–4775.

[39] Friedman, H. S. and Silver, R. C. [2007], *Foundations of health psychology*, Oxford University Press.

[40] Fung, P., Bertero, D., Wan, Y., Dey, A., Chan, R. H. Y., Siddique, F. B., Yang, Y., Wu, C.-S. and Lin, R. [2016], 'Towards empathetic human-robot interactions', *arXiv preprint arXiv:1605.04072* .

[41] Glanz, K., Rimer, B. K. and Viswanath, K. [2008], *Health behavior and health education: theory, research, and practice*, John Wiley & Sons.

[42] Goubert, L., Craig, K. D., Vervoort, T., Morley, S., Sullivan, M., de CAC, W., Cano, A. and Crombez, G. [2005], 'Facing others in pain: the effects of empathy', *Pain* **118**(3), 285–288.

[43] Hartzler, A. and Pratt, W. [2011], 'Managing the personal side of health: how patient expertise differs from the expertise of clinicians', *Journal of medical Internet research* **13**(3).

[44] Heaney, C. A. and Israel, B. A. [2008], 'Social networks and social support', *Health behavior and health education: Theory, research, and practice* **4**, 189–210.

[45] Heckler, C. E., Garland, S. N., Peoples, A. R., Perlis, M. L., Shayne, M., Morrow, G. R., Kamen, C., Hoefler, J. and Roscoe, J. A. [2016], 'Cognitive behavioral therapy for insomnia, but not armodafinil, improves fatigue in cancer survivors with insomnia: a randomized placebo-controlled trial', *Supportive Care in Cancer* **24**(5), 2059–2066.

[46] Hochreiter, S. [1991], 'Untersuchungen zu dynamischen neuronalen netzen', *Diploma, Technische Universität München* p. 91.

[47] Hochreiter, S. and Schmidhuber, J. [1997], 'Long short-term memory', *Neural computation* **9**(8), 1735–1780.

[48] Holt-Lunstad, J., Smith, T. B. and Layton, J. B. [2010], 'Social relationships and mortality risk: a meta-analytic review', *PLoS Med* **7**(7), e1000316.

[49] Holt-Lunstad, J. and Uchino, B. [2015], 'Social support and health', *Health behavior: Theory, research, & practice* pp. 183–204.

[50] House, J. S. [1981], *Work stress and social support*.

[51] Hubel, D. H. and Wiesel, T. N. [1968], 'Receptive fields and functional architecture of monkey striate cortex', *The Journal of physiology* **195**(1), 215–243.

[52] Johnson, R. and Zhang, T. [2014], 'Effective use of word order for text categorization with convolutional neural networks', *arXiv preprint* .

[53] Kardos, P., Leidner, B., Pléh, C., Soltész, P. and Unoka, Z. [2017], 'Empathic people have more friends: Empathic abilities predict social network size and position in social network predicts empathic efforts', *Social Networks* **50**, 1–5.

[54] Katz, P., Singleton, M. and Wicentowski, R. [2007], Swat-mp: the semeval-2007 systems for task 5 and task 14, *in* '4th international workshop on semantic evaluations', pp. 308–313.

[55] Khanpour, H., Caragea, C. and Biyani, P. [2017], Identifying empathetic messages in online health communities, *in* 'Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 2: Short Papers)', Vol. 2, pp. 246–251.

[56] Khanpour, H., Caragea, C. and Biyani, P. [2018], Identifying emotional support in online health communities, *in* 'AAAI press (Volume 2: Short Papers)'.

[57] Khanpour, H., Guntakandla, N. and Nielsen, R. [2016], Dialogue act classification in domain-independent conversations using a deep recurrent neural network, *in* 'Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers', pp. 2012–2021.

[58] Kim, Y. [2014], 'Convolutional neural networks for sentence classification', *arXiv preprint arXiv:1408.5882* .

[59] Kim, Y., Jernite, Y., Sontag, D. and Rush, A. M. [2015], 'Character-aware neural language models', *arXiv preprint arXiv:1508.06615* .

[60] Kim, Y., Jernite, Y., Sontag, D. and Rush, A. M. [2016], Character-aware neural language

models., *in* ‘AAAI’, pp. 2741–2749.

[61] Kroenke, C. H., Quesenberry, C., Kwan, M. L., Sweeney, C., Castillo, A. and Caan, B. J. [2013], ‘Social networks, social support, and burden in relationships, and mortality after breast cancer diagnosis in the life after breast cancer epidemiology (lace) study’, *Breast cancer research and treatment* **137**(1), 261–271.

[62] Lai, S., Xu, L., Liu, K. and Zhao, J. [2015], Recurrent convolutional neural networks for text classification., *in* ‘AAAI’, Vol. 333, pp. 2267–2273.

[63] Lample, G., Ballesteros, M., Subramanian, S., Kawakami, K. and Dyer, C. [2016], ‘Neural architectures for named entity recognition’, *arXiv preprint arXiv:1603.01360* .

[64] Launay, J., Pearce, E., Wlodarski, R., van Duijn, M., Carney, J. and Dunbar, R. I. [2015], ‘Higher-order mentalising and executive functioning’, *Personality and individual differences* **86**, 6–14.

[65] Leite, I., Pereira, A., Mascarenhas, S., Martinho, C., Prada, R. and Paiva, A. [2013], ‘The influence of empathy in human–robot relations’, *International journal of human-computer studies* **71**(3), 250–260.

[66] Litwin, M. S. and McGuigan, K. A. [1999], ‘Accuracy of recall in health-related quality-of-life assessment among men treated for prostate cancer’, *Journal of Clinical Oncology* **17**(9), 2882–2882.

[67] Liu, B. [2012], ‘Sentiment analysis and opinion mining’, *Synthesis lectures on human language technologies* **5**(1), 1–167.

[68] Liu, X., Zhang, S., Wei, F. and Zhou, M. [2011], Recognizing named entities in tweets, *in* ‘Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1’, Association for Computational Linguistics, pp. 359–367.

[69] Magnezi, R., Bergman, Y. S. and Grosberg, D. [2014], ‘Online activity and participation in treatment affects the perceived efficacy of social health networks among patients with chronic illness’, *Journal of medical Internet research* **16**(1).

[70] Malik, S. H. and Coulson, N. S. [2010], ‘Coping with infertility online: an examination of

self-help mechanisms in an online infertility support group', *Patient education and counseling* **81**(2), 315–318.

[71] Mamykina, L., Nakikj, D. and Elhadad, N. [2015], Collective sensemaking in online health forums, *in* 'Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems', ACM, pp. 3217–3226.

[72] Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J. R., Bethard, S. and McClosky, D. [2014], The stanford corenlp natural language processing toolkit., *in* 'ACL (System Demonstrations)', pp. 55–60.

[73] Mayshak, R., Sharman, S. J., Zinkiewicz, L. and Hayley, A. [2017], 'The influence of empathy and self-presentation on engagement with social networking website posts', *Computers in Human Behavior* **71**, 362–377.

[74] Medeiros, L. and Bosse, T. [2016], Empirical analysis of social support provided via social media, *in* 'International Conference on Social Informatics', Springer, pp. 439–453.

[75] Mehra, A., Dixon, A. L., Brass, D. J. and Robertson, B. [2006], 'The social network ties of group leaders: Implications for group performance and leader reputation', *Organization science* **17**(1), 64–79.

[76] Mikolov, T., Karafiát, M., Burget, L., Cernockỳ, J. and Khudanpur, S. [2010], Recurrent neural network based language model., *in* 'Interspeech', Vol. 2, p. 3.

[77] Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S. and Dean, J. [2013], Distributed representations of words and phrases and their compositionality, *in* 'Advances in neural information processing systems', pp. 3111–3119.

[78] Miotto, R., Wang, F., Wang, S., Jiang, X. and Dudley, J. T. [2017], 'Deep learning for healthcare: review, opportunities and challenges', *Briefings in Bioinformatics* p. bbx044.

[79] Mohammad, S. M. [2012], # emotional tweets, *in* 'Sixth International Workshop on Semantic Evaluation', pp. 246–255.

[80] Mohammad, S. M. and Turney, P. D. [2013], 'Crowdsourcing a word–emotion association lexicon', *Computational Intelligence* **29**(3), 436–465.

[81] Moore, T. A. and Wilson, M. E. [2011], 'Feeding intolerance: a concept analysis', *Advances*

in Neonatal Care **11**(3), 149–154.

[82] Nambisan, P. [2011], 'Information seeking and social support in online health communities: impact on patients' perceived empathy', *Journal of the American Medical Informatics Association* **18**(3), 298–304.

[83] Narouei, M., Khanpour, H. and Takabi, H. [2017], Identification of access control policy sentences from natural language policy documents, *in* 'IFIP Annual Conference on Data and Applications Security and Privacy', Springer, pp. 82–100.

[84] Narouei, M., Khanpour, H., Takabi, H., Parde, N. and Nielsen, R. [2017], Towards a top-down policy engineering framework for attribute-based access control, *in* 'Proceedings of the 22nd ACM on Symposium on Access Control Models and Technologies', ACM, pp. 103–114.

[85] Olson, J. K. [1995], 'Relationships between nurse-expressed empathy, patient-perceived empathy and patient distress', *Journal of Nursing Scholarship* **27**(4), 317–322.

[86] Palaz, D., Collobert, R. et al. [2015], Analysis of cnn-based speech recognition system using raw speech as input, Technical report, Idiap.

[87] Pamukçu, B. and Meydan, B. [2010], 'The role of empathic tendency and perceived social support in predicting loneliness levels of college students', *Procedia-Social and Behavioral Sciences* **5**, 905–909.

[88] Pascanu, R., Mikolov, T. and Bengio, Y. [2013], 'On the difficulty of training recurrent neural networks.', *ICML (3)* **28**, 1310–1318.

[89] Qiu, B., Zhao, K., Mitra, P., Wu, D., Caragea, C., Yen, J., Greer, G. E. and Portier, K. [2011], Get online support, feel better–sentiment analysis and dynamics in an online cancer survivor community, *in* 'PASSAT and SocialCom', IEEE, pp. 274–281.

[90] Ravert, R. D., Hancock, M. D. and Ingersoll, G. M. [2004], 'Online forum messages posted by adolescents with type 1 diabetes', *The Diabetes Educator* **30**(5), 827–834.

[91] Redelmeier, D. A. and Kahneman, D. [1996], 'Patients' memories of painful medical treatments: Real-time and retrospective evaluations of two minimally invasive procedures', *Pain* **66**(1), 3–8.

[92] Řehůřek, R. and Sojka, P. [2010], Software Framework for Topic Modelling with Large Cor-

pora, *in* 'LREC', Valletta, Malta, pp. 45–50.

[93] Rogers, L. Q., Courneya, K. S., Anton, P. M., Hopkins-Price, P., Verhulst, S., Robbs, R. S., Vicari, S. K. and McAuley, E. [2016], 'Social cognitive constructs did not mediate the beat cancer intervention effects on objective physical activity behavior based on multivariable path analysis', *Annals of Behavioral Medicine* pp. 1–6.

[94] Rozmovits, L. and Ziebland, S. [2004], 'What do patients with prostate or breast cancer want from an internet site? a qualitative study of information needs', *Patient education and counseling* **53**(1), 57–64.

[95] Rumelhart, D. E., Hinton, G. E. and Williams, R. J. [1986], 'Learning representations by back-propagating errors', *nature* **323**(6088), 533.

[96] Salzer, M. S., Palmer, S. C., Kaplan, K., Brusilovskiy, E., Ten Have, T., Hampshire, M., Metz, J. and Coyne, J. C. [2010], 'A randomized, controlled study of internet peer-to-peer interactions among women newly diagnosed with breast cancer', *Psycho-Oncology* **19**(4), 441–446.

[97] Shanahan, J. G., Qu, Y. and Wiebe, J. [2006], *Computing attitude and affect in text: Theory and applications*, Vol. 20, Springer.

[98] Sheikh, K. [1999], 'Computer-based support systems for women with breast cancer', *JAMA* **281**(14), 1268–1269.

[99] Shennan, A. and Bewley, S. [2005], 'Are virtual communities good for our health?', *birth* **90**, F134–40.

[100] Socher, R., Perelygin, A., Wu, J. Y., Chuang, J., Manning, C. D., Ng, A. Y., Potts, C. et al. [2013], Recursive deep models for semantic compositionality over a sentiment treebank, *in* 'EMNLP', Vol. 1631, Citeseer, p. 1642.

[101] Strapparava, C. and Mihalcea, R. [2007], Semeval-2007 task 14: Affective text, *in* '4th Semantic Evaluations', pp. 70–74.

[102] Strapparava, C. and Mihalcea, R. [2008], Learning to identify emotions in text, *in* 'ACM Applied computing', pp. 1556–1560.

[103] Strapparava, C., Mihalcea, R. and Battocchi, A. [2012], A parallel corpus of music and lyrics annotated with emotions, *in* 'Proceedings of the Eighth International Conference on Lan-

guage Resources and Evaluation, LREC 2012, Istanbul, Turkey, May 23-25, 2012', pp. 2343–2346.

[104] Strapparava, C., Valitutti, A. et al. [2004], Wordnet affect: an affective extension of wordnet., *in* 'LREC', Vol. 4, pp. 1083–1086.

[105] Sykora, M. D., Jackson, T., O'Brien, A. and Elayan, S. [2013], 'Emotive ontology: Extracting fine-grained emotions from terse, informal messages'.

[106] Tang, D., Qin, B. and Liu, T. [2016], 'Aspect level sentiment classification with deep memory network', *arXiv preprint arXiv:1605.08900* .

[107] Tuarob, S., Tucker, C. S., Salathe, M. and Ram, N. [2014], 'An ensemble heterogeneous classification methodology for discovering health-related knowledge in social media messages', *Journal of biomedical informatics* **49**, 255–268.

[108] Valente, T. W. [2010], *Social networks and health: Models, methods, and applications*, Vol. 1, Oxford University Press New York.

[109] van der Eijk, M., Faber, M. J., Aarts, J. W., Kremer, J. A., Munneke, M. and Bloem, B. R. [2013], 'Using online health communities to deliver patient-centered care to people with chronic conditions', *Journal of medical Internet research* **15**(6), e115.

[110] Wang, M., Lu, Z., Li, H., Jiang, W. and Liu, Q. [2015], '*gen* cnn: A convolutional architecture for word sequence prediction', *arXiv preprint arXiv:1503.05034* .

[111] Wang, W. Y. [2017], '" liar, liar pants on fire": A new benchmark dataset for fake news detection', *arXiv preprint arXiv:1705.00648* .

[112] Wang, X., Zhao, K. and Street, I. N. [2014*a*], Predicting user engagement in online health communities based on social support activities, *in* 'Ninth INFORMS Workshop on Data Mining and Analytics'.

[113] Wang, X., Zhao, K. and Street, N. [2014*b*], Social support and user engagement in online health communities, *in* 'Smart Health', Springer, pp. 97–110.

[114] Wang, Y.-C., Kraut, R. and Levine, J. M. [2012], To stay or leave?: the relationship of emotional and informational support to commitment in online health support groups, *in* 'CSCW', ACM, pp. 833–842.

[115] Wilson, T., Wiebe, J. and Hoffmann, P. [2005], Recognizing contextual polarity in phrase-level sentiment analysis, *in* 'Proceedings of the conference on human language technology and empirical methods in natural language processing', Association for Computational Linguistics, pp. 347–354.

[116] Xiao, Y. and Cho, K. [2016], 'Efficient character-level document classification by combining convolution and recurrent layers', *arXiv preprint arXiv:1602.00367* .

[117] Yan, J. L. S. and Turtle, H. R. [2016], Exploring fine-grained emotion detection in tweets, *in* 'Proceedings of NAACL-HLT', pp. 73–80.

[118] Yang, J., Nguyen, M. N., San, P. P., Li, X. and Krishnaswamy, S. [2015], Deep convolutional neural networks on multichannel time series for human activity recognition., *in* 'IJCAI', pp. 3995–4001.

[119] Youyou, W., Kosinski, M. and Stillwell, D. [2015], 'Computer-based personality judgments are more accurate than those made by humans', *Proceedings of the National Academy of Sciences* **112**(4), 1036–1040.

[120] Yu, H. and Hatzivassiloglou, V. [2003], Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences, *in* 'Proceedings of the 2003 conference on Empirical methods in natural language processing', Association for Computational Linguistics, pp. 129–136.

[121] Yu, J. and Kirk, M. [2008], 'Measurement of empathy in nursing research: systematic review', *Journal of Advanced Nursing* **64**(5), 440–454.

[122] Yu, L., Hermann, K. M., Blunsom, P. and Pulman, S. [2014], 'Deep learning for answer sentence selection', *arXiv preprint arXiv:1412.1632* .

[123] Zhang, S., Bantum, E., Owen, J. and Elhadad, N. [2014], Does sustained participation in an online health community affect sentiment?, *in* 'AMIA Annual Symposium Proceedings', Vol. 2014, American Medical Informatics Association, p. 1970.

[124] Zhang, S., Kang, T., Qiu, L., Zhang, W., Yu, Y. and Elhadad, N. [2017], Cataloguing treatments discussed and used in online autism communities, *in* 'Proceedings of the 26th International Conference on World Wide Web', International World Wide Web Conferences Steering

Committee, pp. 123–131.

[125] Zhang, S., OCarroll Bantum, E., Owen, J., Bakken, S. and Elhadad, N. [2017], ‘Online cancer communities as informatics intervention for social support: conceptualization, characterization, and impact’, *Journal of the American Medical Informatics Association* **24**(2), 451–459.

[126] Zhang, Y. and Wallace, B. [2015], ‘A sensitivity analysis of (and practitioners’ guide to) convolutional neural networks for sentence classification’, *arXiv:1510.03820* .

[127] Zhao, K., Yen, J., Greer, G., Qiu, B., Mitra, P. and Portier, K. [2014], ‘Finding influential users of online health communities: a new metric based on sentiment influence’, *Journal of the American Medical Informatics Association* **21**(e2), e212–e218.